EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS

Workbook for Mathematics I

Jana Bělohlávková, Jan Kotůlek

VSB TECHNICAL
UNIVERSITY
OF OSTRAVA

Thanks

This material should provide a guide for the study of the subject Mathematics I at VSB–Technical University of Ostrava, either for Erasmus students or for students of the study programmes taught in English.
It is based on lecture notes of the authors, who have been teaching the subject for the last 10 years. Jana Bělohlávková wrote the chapters on linear algebra, Jan Kotůlek wrote the chapters on calculus.

We thank our colleagues, especially Arnošt Žídek, for many fruitful discussions on mathematics education. Many ideas in the text were inspired by books by Gilbert Strang.

The text was written with the financial support of the project Technology for the Future 2.0, CZ.02.2.69/0.0/0.0/18_058/0010212.

Contents

Workbook for Mathematics I

Linear Algebra

Jana Bělohlávková

1 System of linear equations

Definition

A **system of linear equations** (or **linear system**) is a set of linear equations in the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

The a_{ij} are called the **coefficients** of the system and b_i are called the **right-hand side** of the system.
A **solution** of a linear system is any tuple of numbers that makes each equation a true statement. The set of all solutions of a system is called a **general solution** of the system. Two systems with equal solution sets are called **equivalent**. A system with no solution is called **inconsistent**.

Definition

An array of numbers such as the one below is called a **matrix**. (See page 12 for a formal definition.) The matrix **A** containing the coefficients of a linear system is called the **coefficient matrix** of the system. The matrix **A|b** which contains also the right-hand side is called the **augmented matrix** of the system.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \qquad \mathbf{A}|\mathbf{b} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

Below is an example of a system of three linear equations in four variables x_1, x_2, x_3 and x_4 and the corresponding augmented matrix of the system.

$$\begin{aligned} 3x_1 - x_2 + 2x_3 + x_4 &= 4 \\ -x_1 + x_2 - x_3 + 5x_4 &= -2 \\ x_1 + 0.25x_4 &= 0 \end{aligned} \qquad \begin{pmatrix} 3 & -1 & 2 & 1 & 4 \\ -1 & 1 & -1 & 5 & -2 \\ 1 & 0 & 0 & 0.25 & 0 \end{pmatrix}$$

Some operations with equations do not change the solution of the system. The simplest of these are called elementary operations. If such operations are performed on the rows of the augmented matrix of the system, they are called elementary row operations.

Definition

There are three types of **elementary row operations**:
- swapping two rows,
- multiplying a row by a nonzero number,
- adding a multiple of one row to another row.

1.1 Gaussian elimination

Definition

A matrix is said to be in **row echelon form** if each row except for the first one starts with more zeros than the row above it.

Gaussian elimination is sequence of elementary row operations performed on the augmented matrix of a linear system to convert the matrix into row echelon form. The corresponding equivalent system then is solvable by back substitution.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{17} & b_1 \\ a_{21} & a_{22} & \cdots & a_{27} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{61} & a_{26} & \cdots & a_{67} & b_6 \end{pmatrix} \xrightarrow{\text{el. row oper.}} \cdots \longrightarrow \begin{pmatrix} \circledast & * & * & * & * & * & * & * \\ 0 & \circledast & * & * & * & * & * & * \\ 0 & 0 & 0 & \circledast & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \circledast & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$\circledast$ nonzero pivot
$*$ any number
basic (pivot) columns

augmented matrix of a linear system **row echelon form**

The "shape" of the echelon form of any matrix and in particular the position and number of pivots is uniquely determined by the matrix.

There are three possibilities for the number of solutions of a linear system.

• A system has **no solution** if one of the pivots sits in the right-hand side column.

$$\left(\begin{array}{cccccc|c} \circledast & * & * & * & * & * & * \\ 0 & 0 & \circledast & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \circledast & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \circledast \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

• A system has **one solution** if none of the pivots sit in the right-hand side column and the number of variables is equal to the number of nonzero rows.

$$\left(\begin{array}{cccc|c} \circledast & * & * & * & * \\ 0 & \circledast & * & * & * \\ 0 & 0 & \circledast & * & * \\ 0 & 0 & 0 & \circledast & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

• A system has **infinitely many solutions** if none of the pivots sit in the right-hand side column and the number of variables is less than the number of nonzero rows.

$$\left(\begin{array}{cccccc|c} \circledast & * & * & * & * & * & * \\ 0 & 0 & \circledast & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \circledast & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Definition

The **rank** of a matrix $\mathbf{A}$ is the number of nonzero rows in the matrix in row echelon form obtained from matrix $\mathbf{A}$ by Gaussian elimination. It is denoted by $\text{rank}(\mathbf{A})$.

Example 1

Solve the following system using Gaussian elimination.

$$\begin{aligned} x_1 + x_2 - 7x_3 + 3x_4 - 3x_5 &= 6 \\ 6x_2 + 6x_3 + 6x_4 + 3x_5 &= 24 \\ 2x_1 + 9x_2 - 7x_3 + 11x_4 - 7x_5 &= 23 \\ -3x_1 - x_2 + 24x_3 - 4x_4 + 11x_5 &= 4 \\ 3x_1 + 6x_2 - 17x_3 + 17x_4 - 10x_5 &= 45 \end{aligned}$$

$$\left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & 6 & 6 & 6 & 3 & 24 \\ 2 & 9 & -7 & 11 & -7 & 23 \\ -3 & -1 & 24 & -4 & 11 & 4 \\ 3 & 6 & -17 & 17 & -10 & 45 \end{array}\right) \begin{array}{l} \\ R_2/3 \\ R_3 - 2R_1 \\ R_4+3R_1 \\ R_5-3R_1 \end{array} \to \left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & \textcircled{2} & 2 & 2 & 1 & 8 \\ 0 & 7 & 7 & 5 & -1 & 11 \\ 0 & 2 & 3 & 5 & 2 & 22 \\ 0 & 3 & 4 & 8 & -1 & 27 \end{array}\right) \begin{array}{l} \\ \\ 2R_3 - 7R_2 \\ R_4-R_2 \\ 2R_5-3R_2 \end{array} \to$$

$$\to \left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & \textcircled{2} & 2 & 2 & 1 & 8 \\ 0 & 0 & 0 & -4 & -9 & -34 \\ 0 & 0 & 1 & 3 & 1 & 14 \\ 0 & 0 & 2 & 10 & -5 & 30 \end{array}\right) R_3 \leftrightarrow R_4 \to \left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & \textcircled{2} & 2 & 2 & 1 & 8 \\ 0 & 0 & \textcircled{1} & 3 & 1 & 14 \\ 0 & 0 & 0 & -4 & -9 & -34 \\ 0 & 0 & 2 & 10 & -5 & 30 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ R_5-2R_3 \end{array} \to$$

$$\to \left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & \textcircled{2} & 2 & 2 & 1 & 8 \\ 0 & 0 & \textcircled{1} & 3 & 1 & 14 \\ 0 & 0 & 0 & \textcircled{-4} & -9 & -34 \\ 0 & 0 & 0 & 4 & -7 & 2 \end{array}\right) \begin{array}{l} \\ \\ \\ \\ R_5+R_4 \end{array} \to \left(\begin{array}{ccccc|c} \textcircled{1} & 1 & -7 & 3 & -3 & 6 \\ 0 & \textcircled{2} & 2 & 2 & 1 & 8 \\ 0 & 0 & \textcircled{1} & 3 & 1 & 14 \\ 0 & 0 & 0 & \textcircled{-4} & -9 & -34 \\ 0 & 0 & 0 & 0 & \textcircled{-16} & -32 \end{array}\right)$$

There is no pivot in the right-hand side column, therefore system is consistent. Since the number of variables is equal to the number of nonzero rows the system has one solution.

Back substitution:

$$\begin{aligned} x_1 + x_2 - 7x_3 + 3x_4 - 3x_5 &= 6 \\ 2x_2 + 2x_3 + 2x_4 + x_5 &= 8 \\ x_3 + 3x_4 + x_5 &= 14 \\ -4x_4 - 9x_5 &= -34 \\ -16x_5 &= -32 \end{aligned}$$

$$\begin{aligned} x_1 - 1 - 0 + 3\cdot 4 - 3\cdot 2 = 6 &\longrightarrow x_1 = 1 \\ 2x_2 + 0 + 2\cdot 4 + 2 = 8 &\longrightarrow x_2 = -1 \\ x_3 + 3\cdot 4 + 2 = 14 &\longrightarrow x_3 = 0 \\ -4x_4 - 9\cdot 2 = -34 &\longrightarrow x_4 = 4 \\ &\longrightarrow x_5 = 2 \end{aligned}$$

The solution is:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 4 \\ 2 \end{pmatrix}$$

Verify that the solution is correct:

R_1: $13\cdot 1 + 13\cdot 5 + 13\cdot 6 - 13\cdot 10 = 26$ ✓

R_2: $-1\cdot 1 - 5 - 2\cdot 6 + 2\cdot 10 = 2$ ✓

R_3: $2\cdot 1 + 5\cdot 5 - 3\cdot 6 - 10 = -1$ ✓

R_4: $5\cdot 1 + 7\cdot 5 - 3\cdot 6 - 2\cdot 10 = 2$ ✓

R_5: $13\cdot 1 + 13\cdot 5 + 13\cdot 6 - 13\cdot 10 = 26$ ✓

Example 2

Solve the following system using Gaussian elimination.

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 \quad &= 1 \\ x_1 - 2x_2 + 3x_3 + 2x_4 - x_5 &= 5 \\ x_1 \quad + 5x_3 \quad - x_5 &= 6 \\ 3x_1 - 6x_2 + 6x_3 - 3x_4 \quad &= 3 \\ 2x_2 + 3x_3 + x_4 - x_5 &= 4 \end{aligned}$$

$$\left(\begin{array}{ccccc|c} \textcircled{1} & -2 & 2 & -1 & 0 & 1 \\ 1 & -2 & 3 & 2 & -1 & 5 \\ 1 & 0 & 5 & 0 & -1 & 6 \\ 3 & -6 & 6 & -3 & 0 & 3 \\ 0 & 2 & 3 & 1 & -1 & 4 \end{array}\right) \begin{array}{l} \\ R_2 - R_1 \\ R_3 - R_1 \\ R_4 - 3R_1 \\ \\ \end{array} \rightarrow \left(\begin{array}{ccccc|c} \textcircled{1} & -2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 3 & -1 & 4 \\ 0 & 2 & 3 & 1 & -1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 4 \end{array}\right) \begin{array}{l} \\ R_2 \leftrightarrow R_3 \\ \\ R_4 \leftrightarrow R_5 \\ \\ \end{array} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccccc|c} \textcircled{1} & -2 & 2 & -1 & 0 & 1 \\ 0 & \textcircled{2} & 3 & 1 & -1 & 5 \\ 0 & 0 & 1 & 3 & -1 & 4 \\ 0 & 2 & 3 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \\ \\ R_4 - R_2 \\ \\ \end{array} \rightarrow \left(\begin{array}{ccccc|c} \textcircled{1} & -2 & 2 & -1 & 0 & 1 \\ 0 & \textcircled{2} & 3 & 1 & -1 & 5 \\ 0 & 0 & \textcircled{1} & 3 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

This matrix corresponds to the equations:

$$\begin{aligned} x_1 - 2x_2 + 2x_3 - x_4 \quad &= 1 \\ 2x_2 + 3x_3 + x_4 - x_5 &= 5 \\ x_3 + 3x_4 - x_5 &= 4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= -1 \end{aligned}$$

There are no values such that $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = -1$.
The system has no solution.

Example 3

Solve the following system using Gaussian elimination.

$$\begin{aligned} 2x_1 + x_2 - x_3 \qquad + 8x_5 &= 1 \\ -2x_1 + 2x_2 + 2x_3 + 2x_4 - 4x_5 &= 0 \\ 3x_2 + x_3 + 2x_4 + 5x_5 &= 1 \\ 3x_1 + 6x_2 \qquad + 3x_4 + 17x_5 &= 3 \\ 2x_1 + x_2 - x_3 \qquad + 11x_5 &= 1 \end{aligned}$$

$$\left(\begin{array}{ccccc|c} \textcircled{2} & 1 & -1 & 0 & 8 & 1 \\ -2 & 2 & 2 & 2 & -4 & 0 \\ 0 & 3 & 1 & 2 & 5 & 1 \\ 3 & 6 & 0 & 3 & 17 & 3 \\ 2 & 1 & -1 & 0 & 11 & 1 \end{array}\right) \begin{matrix} \\ R_2+R_1 \\ \\ 2R_4-3R_1 \\ R_5-R_1 \end{matrix} \rightarrow \left(\begin{array}{ccccc|c} \textcircled{2} & 1 & -1 & 0 & 8 & 1 \\ 0 & \textcircled{3} & 1 & 2 & 4 & 1 \\ 0 & 3 & 1 & 2 & 5 & 1 \\ 0 & 9 & 3 & 6 & 10 & 3 \\ 0 & 0 & 0 & 0 & 3 & 0 \end{array}\right) \begin{matrix} \\ \\ R_3-R_2 \\ R_4-3R_2 \\ R_5/3 \end{matrix} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccccc|c} \textcircled{2} & 1 & -1 & 0 & 8 & 1 \\ 0 & \textcircled{3} & 1 & 2 & 4 & 1 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) \begin{matrix} \\ \\ \\ R_4+2R_3 \\ R_4-R_3 \end{matrix} \rightarrow \left(\begin{array}{ccccc|c} \textcircled{2} & 1 & -1 & 0 & 8 & 1 \\ 0 & \textcircled{3} & 1 & 2 & 4 & 1 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

There is no pivot in the right-hand side column, therefore system is consistent. Since the number of variables is less than the number of nonzero rows the system has infinitely many solutions.

The variables in the pivotal columns x_1, x_2 and x_5 are called **basic variables**. The other variables x_3 and x_4 are called **free variables**.

Back substitution:

$$\begin{aligned} 2x_1 + x_2 - x_3 + 8x_5 &= 1 & x_1 + 1/3 - 2t/3 - s/3 - s + 8\cdot 0 &= 1 & \longrightarrow\ & x_1 = 1/3 + t/3 + 2s/3 \\ 3x_2 + x_3 + 2x_4 + 4x_5 &= 1 & 3x_2 + s + 2t + 4\cdot 0 &= 1 & \longrightarrow\ & x_2 = 1/3 - 2t/3 - s/3 \\ & & & & \longrightarrow\ & x_3 = s \\ & & & & \longrightarrow\ & x_4 = t \\ x_5 &= 0 & & & \longrightarrow\ & x_5 = 0 \qquad t, s \in \mathbb{R} \end{aligned}$$

The solution is:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 1/3 \\ -2/3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 2/3 \\ -1/3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Verify that the solution is correct:

R_1: $2(1/3 + t/3 + 2s/3) + (1/3 - 2t/3 - s/3) - s + 0 = 1$ ✓

R_2: $-2(1/3 + t/3 + 2s/3) + 2(1/3 - 2t/3 - s/3) + 2s + 2t = 0$ ✓

R_3: $0 + 3(1/3 - 2t/3 - s/3) + s + 2t + 0 = 1$ ✓

R_4: $3(1/3 + t/3 + 2s/3) + 6(1/3 - 2t/3 - s/3) + 3t + 0 = 3$ ✓

R_5: $2(1/3 + t/3 + 2s/3) + (1/3 - 2t/3 - s/3) - s + 0 = 1$ ✓

1.2 Gauss-Jordan elimination

Definition

A matrix is said to be in **reduced row echelon form** if it is in row echelon form, all pivots are 1 and all entries not only below but also above them are zero.

Gauss-Jordan elimination is sequence of elementary row operations performed on the augmented matrix of a linear system to convert the matrix into reduced row echelon form.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{17} & b_1 \\ a_{21} & a_{22} & \cdots & a_{27} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{61} & a_{26} & \cdots & a_{67} & b_6 \end{pmatrix} \xrightarrow{\text{el. row oper.}} \cdots \longrightarrow \begin{pmatrix} \textcircled{1} & 0 & * & 0 & * & * & 0 & * \\ 0 & \textcircled{1} & * & 0 & * & * & 0 & * \\ 0 & 0 & 0 & \textcircled{1} & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{1} & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$*$ any number
basic (pivot) columns

augmented matrix of a linear system **reduced row echelon form**

There is a unique matrix in row echelon form associated with any matrix **A**. It is usually denoted by rref(**A**).

Example 4

Solve the following system using Gauss-Jordan elimination.

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 3x_4 + 3x_5 &= 0 \\ 2x_1 + 4x_2 - 3x_3 + 10x_4 + 13x_5 &= 8 \\ -3x_1 - 6x_2 + 15x_3 + x_4 + 7x_5 &= 20 \end{aligned}$$

$$\left(\begin{array}{ccccc|c} \textcircled{1} & 2 & -3 & 3 & 3 & 0 \\ 2 & 4 & -3 & 10 & 13 & 8 \\ -3 & -6 & 15 & 1 & 7 & 20 \end{array}\right) \begin{matrix} \\ R_2-2R_1 \longrightarrow \\ R_3+3R_1 \end{matrix} \left(\begin{array}{ccccc|c} \textcircled{1} & 2 & -3 & 3 & 3 & 0 \\ 0 & 0 & \textcircled{3} & 4 & 7 & 8 \\ 0 & 0 & 6 & 10 & 16 & 20 \end{array}\right) \begin{matrix} \\ \longrightarrow \\ R_3-2R_2 \end{matrix} \left(\begin{array}{ccccc|c} \textcircled{1} & 2 & -3 & 3 & 3 & 0 \\ 0 & 0 & \textcircled{3} & 4 & 7 & 8 \\ 0 & 0 & 0 & \textcircled{2} & 2 & 4 \end{array}\right) \begin{matrix} 2R_1-3R_3 \\ R_2-2R_3 \\ \end{matrix}$$

At this point, the augmented matrix is in echelon form. The following additional row operations are performed to transform the matrix to the reduced echelon form.

$$\longrightarrow \left(\begin{array}{ccccc|c} \textcircled{2} & 4 & -6 & 0 & 0 & -12 \\ 0 & 0 & \textcircled{3} & 0 & 3 & 0 \\ 0 & 0 & 0 & \textcircled{2} & 2 & 4 \end{array}\right) \begin{matrix} R_1+2R_2 \\ \longrightarrow \\ \end{matrix} \left(\begin{array}{ccccc|c} \textcircled{2} & 4 & 0 & 0 & 6 & -12 \\ 0 & 0 & \textcircled{3} & 0 & 3 & 0 \\ 0 & 0 & 0 & \textcircled{2} & 2 & 4 \end{array}\right) \begin{matrix} R_1/2 \\ R_2/3 \longrightarrow \\ R_3/2 \end{matrix} \left(\begin{array}{ccccc|c} \textcircled{1} & 2 & 0 & 0 & 3 & -6 \\ 0 & 0 & \textcircled{1} & 0 & 1 & 0 \\ 0 & 0 & 0 & \textcircled{1} & 1 & 2 \end{array}\right)$$

The solution is:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -3 \\ 0 \\ -1 \\ -1 \\ 1 \end{pmatrix} + s \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Verify that the solution is correct:

R_1: $-6 - 3t - 2s + 2s + 3t + 3(2-t) + 3t = 0$ ✓

R_2: $2(-6-3t-2s) + 4s + 3t + 10(2-t) + 13t = 8$ ✓

R_3: $-3(-6-3t-2s) - 6s - 15t + 2 - t + 7t = 20$ ✓

1.3 Homogeneous and nonhomogeneous systems

Definition

A system of linear equations with the right-hand side consisting entirely of zeros is said to be **homogeneous**.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

A system with at least one nonzero number on the right-hand side is called **nonhomogeneous**.

Regardless of the value of the coefficients a homogeneous system has always at least one solution, the **trivial solution** consisting of all zeros.
There is a close relation between the solution of a nonhomogeneous system and the solution of the associated homogeneous one.

Example 5

Solve the following systems.

$$\begin{aligned} x_1 + 4x_2 + 6x_3 + 33x_4 + 2x_5 &= 27 \\ x_1 + 4x_2 + 8x_3 + 43x_4 + 2x_5 &= 31 \\ x_1 + 4x_2 + 2x_3 + 13x_4 + x_5 &= 12 \end{aligned} \qquad \begin{aligned} x_1 + 4x_2 + 6x_3 + 33x_4 + 2x_5 &= 0 \\ x_1 + 4x_2 + 8x_3 + 43x_4 + 2x_5 &= 0 \\ x_1 + 4x_2 + 2x_3 + 13x_4 + x_5 &= 0 \end{aligned}$$

a nonhomogeneous system — the associated homogenous system

The elimination can be performed on both systems at the same time.

$$\left(\begin{array}{ccccc|cc} 1 & 4 & 6 & 33 & 2 & 27 & 0 \\ 1 & 4 & 8 & 43 & 2 & 31 & 0 \\ 1 & 4 & 2 & 13 & 1 & 12 & 0 \end{array}\right) \begin{array}{l} \\ R_2 - R_1 \\ R_3 - R_1 \end{array} \rightarrow \left(\begin{array}{ccccc|cc} 1 & 4 & 6 & 33 & 2 & 27 & 0 \\ 0 & 0 & 2 & 10 & 0 & 4 & 0 \\ 0 & 0 & -4 & -20 & -1 & -15 & 0 \end{array}\right) \begin{array}{l} R_1 - 3R_2 \\ \\ R_3 + 2R_2 \end{array} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccccc|cc} 1 & 4 & 0 & 3 & 2 & 15 & 0 \\ 0 & 0 & 2 & 10 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & -1 & -7 & 0 \end{array}\right) \begin{array}{l} R_1 + 2R_3 \\ R_2/2 \\ R_3/-1 \end{array} \rightarrow \left(\begin{array}{ccccc|cc} 1 & 4 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 5 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 7 & 0 \end{array}\right)$$

$$\begin{aligned} x_1 + 4x_2 + 3x_4 &= 1 \\ \\ x_3 + 5x_4 &= 2 \\ \\ x_5 &= 7 \end{aligned} \qquad \begin{aligned} x_1 &= 1 - 3 - 4st \\ x_2 &= s \\ x_3 &= 2 - 5t \\ x_4 &= t \\ x_5 &= 7 \end{aligned} \qquad \begin{aligned} x_1 + 4x_2 + 3x_4 &= 0 \\ \\ x_3 + 5x_4 &= 0 \\ \\ x_5 &= 0 \end{aligned} \qquad \begin{aligned} x_1 &= -3t - 4s \\ x_2 &= s \\ x_3 &= -5t \\ x_4 &= t \\ x_5 &= 0 \end{aligned}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 7 \end{pmatrix} + t \begin{pmatrix} -3 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = t \begin{pmatrix} -3 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

A general solution of the nonhomogeneous system is the sum of a so called **particular solution** (the green part) and the solution of the associated homogeneous system (the blue part).

2 Matrix algebra

Definition

An array of numbers (real or complex) is called a **matrix**.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

The number a_{ij} in the ith row and jth column is called an **entry** of the matrix. It can be also denoted $(\mathbf{A})_{ij}$.
The **size** of the matrix is denoted $m \times n$ (pronounced "m by n").
The entries $a_{11}, a_{22}, a_{33}, \ldots$ make up the **main diagonal**.

Definition

A matrix is called
- **a square matrix** when it has the same number of rows and columns.
- **a rectangular matrix** when it doesn't have the same number of rows and columns.
- **a zero matrix** when all of its entries are zero. Zero matrices are denoted **O**.

For example:

$$\begin{pmatrix} 1 & 5 & 7 \\ 1 & 4 & 3 \\ 3 & 7 & 8 \end{pmatrix}$$

3×3 square matrix
blue main diagonal

$$\begin{pmatrix} 2 & 4 & 5 & 1 & 1 \\ 0 & 4 & 1 & 4 & 0 \\ 1 & 0 & 5 & 2 & 4 \end{pmatrix}$$

3×5 rectangular matrix
blue main diagonal

$$\begin{pmatrix} 2 & 4 & 5 & 1 \\ 0 & 4 & 1 & 4 \\ 1 & 0 & 5 & 2 \\ 1 & 1 & 5 & 1 \\ 1 & 4 & 0 & 3 \\ 1 & 2 & 0 & 2 \end{pmatrix}$$

6×4 rectangular matrix
blue main diagonal

$$\mathbf{O} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

3×2 zero matrix

Definition

A square matrix is called
- **diagonal** if all entries below and above the main diagonal are zero.
- **an identity matrix** when it has ones on the main diagonal and zeros everywhere else. Identity matrices are denoted **I**.
- **lower triangular** if all the entries above the main diagonal are zero.
- **upper triangular** if all the entries below the main diagonal are zero.
- **symetric** if every entry a_{ij} is equal the entry a_{ji}.

For example:

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

a diagonal matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

an identity matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 2/7 & 8 \end{pmatrix}$$

a lower triangular matrix

$$\begin{pmatrix} 1 & 0 & 1 & 4 \\ 0 & 6 & 1 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

an upper triangular matrix

$$\begin{pmatrix} 2 & 0 & 5 & 3 \\ 0 & -4 & 0 & 1 \\ 5 & 0 & 0 & 8 \\ 3 & 1 & 8 & 7 \end{pmatrix}$$

a symetric matrix

2.1 Matrix operation

Definition

The **product of a number** k **and a matrix** $\mathbf{A}$ is defined to be the matrix obtained by multiplying each entry of $\mathbf{A}$ by k.

$$(k \cdot \mathbf{A})_{ij} = k \cdot a_{ij}$$

For example:

$$3 \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 8 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 0 \\ 9 & 24 \end{pmatrix}$$

Definition

The **sum of two matrices** $\mathbf{A}_{m\times n}$ and $\mathbf{B}_{m\times n}$ is defined to be the $m\times n$ matrix obtained by adding corresponding entries.

$$(\mathbf{A}+\mathbf{B})_{ij} = a_{ij} + b_{ij}$$

For example:

$$\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 8 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 3 & 5 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 5 & 5 \\ 6 & 12 \end{pmatrix}$$

Definition

The **product of two matrices** $\mathbf{A}_{m\times p}$ and $\mathbf{B}_{p\times n}$ is defined to be the $m\times n$ matrix whose ijth entry is obtained by "multiplying" ith row of $\mathbf{A}$ with jth column of $\mathbf{B}$ as follows:

$$(\mathbf{A}\cdot\mathbf{B})_{ij} = \sum_{k=1}^{p} a_{ik} \cdot b_{kj}$$

For example:

a) $$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 7 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1\cdot 4 + 2\cdot 4 + 3\cdot 4 \\ 0\cdot 4 + 7\cdot 4 + 0\cdot 4 \end{pmatrix} = \begin{pmatrix} 24 \\ 28 \end{pmatrix}$$

b) $$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 7 & 0 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 4 & 6 \\ 4 & 7 \end{pmatrix} = \begin{pmatrix} 24 & 38 \\ 38 & 42 \end{pmatrix}$$

Definition

The **transpose** of $\mathbf{A}_{m\times n}$ is defined to be the $n\times m$ matrix $\mathbf{A}^{\mathsf{T}}$ obtained by flipping $\mathbf{A}$ over its main diagonal.

$$(\mathbf{A}^{\mathsf{T}})_{ij} = a_{ji}$$

For example:

$$\begin{pmatrix} 1 & 3 & 7 \\ 2 & 8 & 0 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & 2 \\ 3 & 8 \\ 7 & 0 \end{pmatrix}$$

For matrices **A**, **B** and **C** of the right size, the following properties hold.

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$
$$\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$$
$$k\cdot(\mathbf{A}+\mathbf{B})=k\cdot\mathbf{A}+k\cdot\mathbf{B}$$

$$\mathbf{A}\cdot\mathbf{B}\neq\mathbf{B}\cdot\mathbf{A}$$
$$\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})=(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}$$
$$(\mathbf{A}+\mathbf{B})\cdot\mathbf{C}=\mathbf{A}\cdot\mathbf{C}+\mathbf{B}\cdot\mathbf{C}$$
$$\mathbf{C}\cdot(\mathbf{A}+\mathbf{B})=\mathbf{C}\cdot\mathbf{A}+\mathbf{C}\cdot\mathbf{B}$$

Because $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})=(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}$, there is no need to write parentheses and we can simply write $\mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C}$.

Definition

For any positive integer k **the** k**th power of the square matrix** $\mathbf{A}$ is defined as the product of k matrices **A**.

$$\mathbf{A}^k=\underbrace{\mathbf{A}\cdot\mathbf{A}\cdot\mathbf{A}\ldots\mathbf{A}}_{k\text{ times}}$$

A matrix to the *zero*th power is defined to be the identity matrix of the same size $\mathbf{A}^0=\mathbf{I}$.

Exercise 6

a) Find $3\cdot\mathbf{A}+2\cdot\mathbf{B}^{\mathsf{T}}$ for the following matrices.

i) $\mathbf{A}=\begin{pmatrix}3 & 5\\ 7 & -1\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}1 & 10\\ 4 & -2\end{pmatrix}$ ii) $\mathbf{A}=\begin{pmatrix}2 & -2 & 1\\ 0 & 7 & -1\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}1 & 0\\ -1 & -2\\ 7 & 3\end{pmatrix}$

b) Evaluate the following:

i) $\begin{pmatrix}2 & 1 & 2 & -6 & -5\\ 0 & 1 & -4 & 1 & 0\\ 3 & 1 & 1 & 1 & -3\end{pmatrix}\begin{pmatrix}2\\ 1\\ 0\\ 1\\ 1\end{pmatrix}$ ii) $\begin{pmatrix}3 & 2\\ 0 & 6\\ 5 & 1\\ 4 & -1\\ 1 & 2\end{pmatrix}\begin{pmatrix}2\\ 1\end{pmatrix}$

c) Find $\mathbf{A}\cdot\mathbf{B}$ and $\mathbf{B}\cdot\mathbf{A}$ for the following matrices.

i) $\mathbf{A}=\begin{pmatrix}-2 & 6\\ 1 & -3\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}1 & 2 & 0\\ 2 & 4 & 1\end{pmatrix}$ ii) $\mathbf{A}=\begin{pmatrix}1 & 7 & 3\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}2\\ 0\\ 1\end{pmatrix}$

iii) $\mathbf{A}=\begin{pmatrix}-2 & 6\\ 1 & -3\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$ iv) $\mathbf{A}=\begin{pmatrix}2 & 1\\ 4 & 2\end{pmatrix}$ $\mathbf{B}=\begin{pmatrix}2 & 3\\ -4 & -6\end{pmatrix}$

d) Find $\mathbf{A}^3$ for the following matrices.

i) $\mathbf{A}=\begin{pmatrix}-2 & 6\\ 1 & -3\end{pmatrix}$ ii) $\mathbf{A}=\begin{pmatrix}1 & 2 & 0\\ 1 & 0 & 3\\ 1 & 2 & 1\end{pmatrix}$

2.2 Matrix Inverse

Definition

A square matrix $\mathbf{A}$ is called **invertible** if there exists a matrix $\mathbf{A}^{-1}$ such that

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$$

The matrix $\mathbf{A}^{-1}$ is called the **inverse** of $\mathbf{A}$.
A square matrix which is not invertible is called **singular**.

Although not all matrices are invertible, when an inverse exists, it is unique.
Gauss-Jordan elimination can be used to compute the inverse.

Example 7

Find the inverse of the matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix}$.

$$(\,\mathbf{A} \mid \mathbf{I}\,) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 5 & 6 & 2 & 0 & 1 & 0 \\ 2 & 5 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ R_2-5R_1 \\ R_3-2R_1 \end{array} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -5 & 1 & 0 \\ 0 & 3 & -1 & -32 & 0 & 1 \end{array}\right)_{R_3-3R_2} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -5 & 1 & 0 \\ 0 & 0 & 8 & 13 & -3 & 1 \end{array}\right) \begin{array}{l} 8R_1-R_3 \\ 8R_2+3R_3 \\ \end{array} \rightarrow \left(\begin{array}{ccc|ccc} 8 & 8 & 0 & -5 & 3 & -1 \\ 0 & 8 & 0 & -1 & -1 & 3 \\ 0 & 0 & 8 & 13 & -3 & 1 \end{array}\right)^{R_1-R_2} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 8 & 0 & 0 & -4 & 4 & -4 \\ 0 & 8 & 0 & -1 & -1 & 3 \\ 0 & 0 & 8 & 13 & -3 & 1 \end{array}\right) \begin{array}{l} R_1/8 \\ R_2/8 \\ R_3/8 \end{array} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4/8 & 4/8 & -4/8 \\ 0 & 1 & 0 & -1/8 & -1/8 & 3/8 \\ 0 & 0 & 1 & 13/8 & -3/8 & 1/8 \end{array}\right) = (\,\mathbf{I} \mid \mathbf{A}^{-1}\,)$$

$$\mathbf{A}^{-1} = 1/8 \begin{pmatrix} -4 & 4 & -4 \\ -1 & -1 & 3 \\ 13 & -3 & 1 \end{pmatrix}$$

Verify that the solution is correct:

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix} 1/8 \begin{pmatrix} -4 & 4 & -4 \\ -1 & -1 & 3 \\ 13 & -3 & 1 \end{pmatrix} = 1/8 \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A}^{-1} \cdot \mathbf{A} = 1/8 \begin{pmatrix} -4 & 4 & -4 \\ -1 & -1 & 3 \\ 13 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix} = 1/8 \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example 8

Find the inverse of the matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 3 & 4 & 0 \end{pmatrix}$

$$(\,\mathbf{A}\mid\mathbf{I}\,) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 5 & 6 & 2 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 & 0 & 1 \end{array}\right)_{\substack{R_2-5R_1 \\ R_3-3R_1}} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -5 & 1 & 0 \\ 0 & 1 & -3 & -3 & 0 & 1 \end{array}\right)_{R_3-R_2} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -5 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 & 1 \end{array}\right)$$

The matrix **A** cannot be reduced to an identity matrix because a zero row emerged during elimination. Therefore **A** is singular.

A $n \times n$ matrix is invertible if and only if its rank is n.

For two invertible matrices **A** and **B**, the following properties hold.

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- The product $\mathbf{A} \cdot \mathbf{B}$ is also invertible.
- $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$ (the reverse order law for inversion)

Exercise 9

Find the inverse of the following matrices.

a) $\begin{pmatrix} 6 & 1 & 2 \\ 0 & 3 & -1 \\ 4 & 2 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 1 & -1 \\ 2 & 5 & 8 \\ 2 & 7 & 6 \end{pmatrix}$ c) $\begin{pmatrix} -4 & 18 & 10 \\ -1 & -3 & 0 \\ -1 & 17 & 10 \end{pmatrix}$

d) $\begin{pmatrix} 3 & 5 \\ -1 & -3 \end{pmatrix}$ e) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 4 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}$ f) $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}$

2.3 Matrix equations

A system of linear equations can be written as a matrix equation.

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 5x_1 + 6x_2 + 2x_3 &= 3 \\ 2x_1 + 5x_2 + x_3 &= 2 \end{aligned} \qquad \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}}_{\mathbf{b}}$$

The matrix equations $\mathbf{A} \cdot \mathbf{X} = \mathbf{B}$ and $\mathbf{X} \cdot \mathbf{A} = \mathbf{B}$ can be solved if the matrix $\mathbf{A}$ is invertible:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{X} &= \mathbf{B} \\ \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{X} &= \mathbf{A}^{-1} \cdot \mathbf{B} \\ \underbrace{\mathbf{A}^{-1} \cdot \mathbf{A}}_{\mathbf{I}} \cdot \mathbf{X} &= \mathbf{A}^{-1} \cdot \mathbf{B} \\ \mathbf{I} \cdot \mathbf{X} &= \mathbf{A}^{-1} \cdot \mathbf{B} \\ \mathbf{X} &= \mathbf{A}^{-1} \cdot \mathbf{B} \end{aligned} \qquad \begin{aligned} \mathbf{X} \cdot \mathbf{A} &= \mathbf{B} \\ \mathbf{X} \cdot \mathbf{A} \cdot \mathbf{A}^{-1} &= \mathbf{B} \cdot \mathbf{A}^{-1} \\ \mathbf{X} \cdot \underbrace{\mathbf{A} \cdot \mathbf{A}^{-1}}_{\mathbf{I}} &= \mathbf{B} \cdot \mathbf{A}^{-1} \\ \mathbf{X} \cdot \mathbf{I} &= \mathbf{B} \cdot \mathbf{A}^{-1} \\ \mathbf{X} &= \mathbf{B} \cdot \mathbf{A}^{-1} \end{aligned}$$

Example 10

Solve the following matrix equation $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$.

$$\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}$$

THe matrix $\mathbf{A}$ has the inverse $\mathbf{A}^{-1} = 1/8 \begin{pmatrix} -4 & 4 & -4 \\ -1 & -1 & 3 \\ 13 & -3 & 1 \end{pmatrix}$, see page 15.

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= \mathbf{b} \\ \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} &= \mathbf{A}^{-1} \cdot \mathbf{b} \\ \mathbf{x} &= \mathbf{A}^{-1} \cdot \mathbf{b} \end{aligned}$$

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} = 1/8 \begin{pmatrix} -4 & 4 & -4 \\ -1 & -1 & 3 \\ 13 & -3 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} = 1/8 \begin{pmatrix} -8 \\ 0 \\ 32 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$$

Verify that the solution is correct:

$$\mathbf{A} \cdot \mathbf{x} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 2 & 5 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} = \mathbf{b} \quad \checkmark$$

Example 11

Solve the following matrix equation for the unknown matrix **X**.

$$\mathbf{F}\cdot\mathbf{G}\cdot\mathbf{X}=\mathbf{B} \qquad \mathbf{F}=\begin{pmatrix}0&1\\1&2\end{pmatrix} \quad \mathbf{G}=\begin{pmatrix}2&-5\\-1&4\end{pmatrix} \quad \mathbf{B}=\begin{pmatrix}4&1\\6&3\end{pmatrix}$$

$$(\,\mathbf{F}\mid\mathbf{I}\,)=\left(\begin{array}{cc|cc}0&1&1&0\\1&2&0&1\end{array}\right)\underset{R_2\leftrightarrow R_1}{\longrightarrow}\left(\begin{array}{cc|cc}1&2&0&1\\0&1&1&0\end{array}\right)\overset{R_1-2R_2}{\longrightarrow}\left(\begin{array}{cc|cc}1&0&-2&1\\0&1&1&0\end{array}\right)=(\,\mathbf{I}\mid\mathbf{F}^{-1}\,)$$

$$(\,\mathbf{G}\mid\mathbf{I}\,)=\left(\begin{array}{cc|cc}2&-5&1&0\\-1&4&0&1\end{array}\right)\underset{2R_2+R_1}{\longrightarrow}\left(\begin{array}{cc|cc}2&-5&1&0\\0&3&1&2\end{array}\right)\overset{3R_1+5R_2}{\longrightarrow}\left(\begin{array}{cc|cc}6&0&8&10\\0&3&1&2\end{array}\right)\underset{R_2/3}{\overset{R_1/6}{\longrightarrow}}$$

$$\longrightarrow\left(\begin{array}{cc|cc}1&0&4/3&5/3\\0&1&1/3&2/3\end{array}\right)=(\,\mathbf{I}\mid\mathbf{G}^{-1}\,)$$

The matrices **F** and **G** are invertible. Their inverses are $\mathbf{F}^{-1}=\begin{pmatrix}-2&1\\1&0\end{pmatrix}$ and $\mathbf{G}^{-1}=1/3\begin{pmatrix}4&5\\1&2\end{pmatrix}$.

$$\begin{aligned}\mathbf{F}\cdot\mathbf{G}\cdot\mathbf{X}&=\mathbf{B}\\ \mathbf{F}^{-1}\cdot\mathbf{F}\cdot\mathbf{G}\cdot\mathbf{X}&=\mathbf{F}^{-1}\cdot\mathbf{B}\\ \mathbf{G}\cdot\mathbf{X}&=\mathbf{F}^{-1}\cdot\mathbf{B}\\ \mathbf{G}^{-1}\cdot\mathbf{G}\cdot\mathbf{X}&=\mathbf{G}^{-1}\cdot\mathbf{F}^{-1}\cdot\mathbf{B}\\ \mathbf{X}&=\mathbf{G}^{-1}\cdot\mathbf{F}^{-1}\cdot\mathbf{B}\end{aligned}$$

$$\mathbf{X}=1/3\begin{pmatrix}4&5\\1&2\end{pmatrix}\begin{pmatrix}-2&1\\1&0\end{pmatrix}\begin{pmatrix}4&1\\6&3\end{pmatrix}=$$
$$=1/3\begin{pmatrix}-3&4\\0&1\end{pmatrix}\begin{pmatrix}4&1\\6&3\end{pmatrix}=\begin{pmatrix}4&3\\2&1\end{pmatrix}$$

Verify that the solution is correct:

$$\mathbf{F}\cdot\mathbf{G}\cdot\mathbf{X}=\begin{pmatrix}0&1\\1&2\end{pmatrix}\begin{pmatrix}2&-5\\-1&4\end{pmatrix}\begin{pmatrix}4&3\\2&1\end{pmatrix}=\begin{pmatrix}-1&4\\0&3\end{pmatrix}\begin{pmatrix}4&3\\2&1\end{pmatrix}=\begin{pmatrix}4&1\\6&3\end{pmatrix}=\mathbf{B}\ \checkmark$$

Example 12

Solve the following matrix equation for the unknown matrix **X**.

a) $\mathbf{F}\cdot\mathbf{X}\cdot\mathbf{G}=\mathbf{B}$ $\qquad \mathbf{F}=\begin{pmatrix}0&-1\\2&5\end{pmatrix} \quad \mathbf{G}=\begin{pmatrix}2&-7\\-1&4\end{pmatrix} \quad \mathbf{B}=\begin{pmatrix}-1&8\\-1&12\end{pmatrix}$

b) $\mathbf{F}\cdot\mathbf{G}\cdot\mathbf{X}=\mathbf{B}$ $\qquad \mathbf{F}=\begin{pmatrix}1&0&1\\2&2&3\\5&-1&2\end{pmatrix} \quad \mathbf{G}=\begin{pmatrix}1&0&0\\-3&1&0\\5&0&2\end{pmatrix} \quad \mathbf{B}=\begin{pmatrix}6&2\\9&10\\19&2\end{pmatrix}$

c) $\mathbf{A}\cdot\mathbf{X}=\mathbf{b}$

d) $\mathbf{A}\cdot\mathbf{X}=\mathbf{c}$ $\qquad \mathbf{A}=\begin{pmatrix}1&2&4\\1&3&5\\1&3&6\end{pmatrix} \quad \mathbf{b}=\begin{pmatrix}4\\4\\1\end{pmatrix} \quad \mathbf{c}=\begin{pmatrix}10\\10\\20\end{pmatrix} \quad \mathbf{D}=\begin{pmatrix}0&1\\0&1\\0&1\end{pmatrix}$

e) $\mathbf{A}\cdot\mathbf{X}=\mathbf{D}$

2.4 Elementary matrices

Definition

Elementary matrices are square matrices that can be obtained from the identity matrix by performing one single elementary row operation.

For every elementary row operation there is a elementary matrix such that multiplying by it from the left performs the operation.
For example:

- Swapping two rows:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad \cdots$$

- Multiplying a row by a nonzero number α:

$$\mathbf{E}_{11} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_{22} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_{33} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

- Adding a multiple of one row to another row.

$$\mathbf{E}_{13} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} \quad \mathbf{E}_{12} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \alpha & 1 \end{pmatrix} \quad \cdots$$

Definition

A matrix that can be obtained from the identity matrix by swapping two or more rows is called a **permutation matrix**.

Exercise 13

a) What matrix will swap the first and the third row?

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & 3 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 3 \\ 0 & 2 & 1 \\ 1 & 2 & 0 \\ 4 & 2 & 1 \end{pmatrix}$$

b) What matrix will multiply the third row by five?

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & 3 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ -10 & 10 & 15 \\ 4 & 2 & 1 \end{pmatrix}$$

c) What matrix will add the first row to the third one?

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & 3 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 4 & 3 \\ 4 & 2 & 1 \end{pmatrix}$$

2.5 The LU factorization

Some square matrices **A** can be decomposed into two matrices, a lower triangular matrix **L** and an upper triangular matrix **U** such that $\mathbf{A} = \mathbf{L}\cdot\mathbf{U}$. Such a decomposition is called the **LU factorization** of **A**. It may be found by performing Gaussian elimination. This will be demonstrated on the following example.

Example 14

Find the **LU** factors the matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix}$.

The elimination performed on the matrix **A** produces the matrix **U**.

$$\underset{\mathbf{A}}{\begin{pmatrix} \textcircled{1} & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix}} \xrightarrow{R_2 - 5R_1} \begin{pmatrix} \textcircled{1} & 1 & 1 \\ 0 & 1 & 4 \\ 3 & 5 & 13 \end{pmatrix} \xrightarrow{R_3 - 3R_1} \begin{pmatrix} \textcircled{1} & 1 & 1 \\ 0 & \textcircled{1} & 4 \\ 0 & 2 & 10 \end{pmatrix} \xrightarrow{R_3 - 2R_2} \underset{\mathbf{U}}{\begin{pmatrix} \textcircled{1} & 1 & 1 \\ 0 & \textcircled{1} & 4 \\ 0 & 0 & \textcircled{2} \end{pmatrix}}$$

Each one of these row operations can be carried out as multiplication by an elementary matrix.

$$\underset{\mathbf{E}_{23}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}} \cdot \underset{\mathbf{E}_{13}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}} \cdot \underset{\mathbf{E}_{12}}{\begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \cdot \underset{\mathbf{A}}{\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix}} = \underset{\mathbf{U}}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}}$$

$$\begin{aligned} \mathbf{E}_{23}\cdot\mathbf{E}_{13}\cdot\mathbf{E}_{12}\cdot\mathbf{A} &= \mathbf{U} \\ \mathbf{E}_{13}\cdot\mathbf{E}_{12}\cdot\mathbf{A} &= \mathbf{E}_{23}^{-1}\cdot\mathbf{U} \\ \mathbf{E}_{12}\cdot\mathbf{A} &= \mathbf{E}_{13}^{-1}\cdot\mathbf{E}_{23}^{-1}\cdot\mathbf{U} \\ \mathbf{A} &= \underbrace{\mathbf{E}_{12}^{-1}\cdot\mathbf{E}_{13}^{-1}\cdot\mathbf{E}_{23}^{-1}}_{\mathbf{L}}\cdot\mathbf{U} \end{aligned}$$

$$\mathbf{L} = \underset{\mathbf{E}_{12}^{-1}}{\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \cdot \underset{\mathbf{E}_{13}^{-1}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}} \cdot \underset{\mathbf{E}_{23}^{-1}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}} = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

Thus

$$\underset{\mathbf{A}}{\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix}} = \underset{\mathbf{L}}{\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}} \cdot \underset{\mathbf{U}}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}}$$

The matrix **L** has ones on its diagonal. Entries below the diagonal are called **multipliers**. The multiplier ℓ_{ij} is the number used in the elimination to annihilate the ij position: $R_i - \ell_{ij}R_j$

$$\ell_{ij} = \frac{\text{entry to eliminate in row } i}{\text{pivot in row } j} \qquad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

A square matrix **A** can be decomposed into $\mathbf{L}\cdot\mathbf{U}$ if there was no need to exchange rows during the elimination. Otherwise the factorization has the form $\mathbf{P}\cdot\mathbf{A} = \mathbf{L}\cdot\mathbf{U}$, where **P** is a permutation matrix.

LU factorization is a very useful tool for solving multiple systems with the same coefficient matrix and different right-hand sides.

Example 15

Use LU factorization to solve $\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$

Once the LU factors of $\mathbf{A}$ are known (see example on page 20), it is easy to solve $\mathbf{A}\cdot\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= \mathbf{b} \\ \mathbf{L} \cdot \mathbf{U} \cdot \mathbf{x} &= \mathbf{b} \\ \mathbf{L} \cdot (\underbrace{\mathbf{U} \cdot \mathbf{x}}_{\mathbf{y}}) &= \mathbf{b} \end{aligned}$$

Insted of solving $\mathbf{A}\cdot\mathbf{x} = \mathbf{b}$, we solve two triangular systems $\mathbf{L}\cdot\mathbf{y} = \mathbf{b}$ and $\mathbf{U}\cdot\mathbf{x} = \mathbf{y}$:

First $\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}}_{\mathbf{L}} \cdot \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}}_{\mathbf{b}}$ and then $\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}}_{\mathbf{U}} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}}_{\mathbf{y}}$

Forward substitution to get $\mathbf{y}$:

$$\begin{aligned} y_1 &= 0 \\ y_2 &= 1 - 5y_1 = 1 \\ y_3 &= 4 - 3y_1 - 2y_2 = 2 \end{aligned}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Backward substitution to get $\mathbf{x}$.

$$\begin{aligned} x_1 &= 0 - x_2 - x_3 = 2 \\ x_2 &= 1 - 4x_3 = -3 \\ x_3 &= 2/2 = 1 \end{aligned}$$

Exercise 16

Use LU factorization to solve the systems below.

a) $\begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 9 \\ 3 & 5 & 13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 27 \\ 15 \end{pmatrix}$

b) $\begin{pmatrix} 5 & -15 & 1 \\ -20 & 44 & -3 \\ 10 & 50 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -15 \\ 40 \\ 90 \end{pmatrix}$

c) $\begin{pmatrix} 5 & -15 & 1 \\ -20 & 44 & -3 \\ 10 & 50 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

d) $\begin{pmatrix} 5 & -15 & 1 \\ -20 & 44 & -3 \\ 10 & 50 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -9 \\ 21 \\ 56 \end{pmatrix}$

2.6 Determinants

Definition

The **determinant** of a square 1×1 matrix $\mathbf{A} = (a_{11})$ is defined to be the number a_{11}.

The **determinant** of a square matrix $\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ is defined to be the number

$$\det(\mathbf{A}) = \sum_{k=1}^{n} a_{1k}(-1)^{1+k} M_{1k}$$

where M_{ij} is the determinant of the $(n-1)\times(n-1)$ matrix that results from $\mathbf{A}$ by removing its ith row and its jth column.
The number $C_{ij} = (-1)^{i+j} M_{ij}$ is called the **cofactor** associated with the position ij.
The determinant of the matrix $\mathbf{A}$ is denoted $\det(\mathbf{A})$, $\det \mathbf{A}$, or $|\mathbf{A}|$.

Example 17

Compute the determinant of the matrix a) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 7 & 4 \end{pmatrix}$ b) $\mathbf{A} = \begin{pmatrix} 2 & 4 & 1 \\ 1 & 5 & 2 \\ 3 & 0 & 0 \end{pmatrix}$

a) $\det(\mathbf{A}) = \begin{vmatrix} 1 & 2 \\ 7 & 4 \end{vmatrix} = 1\cdot 4 - 7\cdot 2 = 4 - 14 = -10$

b) $\det(\mathbf{A}) = \begin{vmatrix} 2 & 4 & 1 \\ 1 & 5 & 2 \\ 3 & 0 & 0 \end{vmatrix} = 2\cdot(-1)^2 \cdot \begin{vmatrix} 5 & 2 \\ 0 & 0 \end{vmatrix} + 4\cdot(-1)^3 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} + 1\cdot(-1)^4 \cdot \begin{vmatrix} 1 & 5 \\ 3 & 0 \end{vmatrix}$

$= 2\cdot 0 - 4\cdot(-6) + 1\cdot(-15) = 9$

The rule of Sarrus for the determinant of a 3×3 matrix $\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

$$\det(\mathbf{A}) = \left|\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right. = aei + bfg + cdh - ceg - afh - bdi$$

Example 18

Compute the determinant of the matrix $\mathbf{A} = \begin{pmatrix} 2 & 4 & 1 \\ 1 & 5 & 2 \\ 3 & 0 & 0 \end{pmatrix}$.

$$\det(\mathbf{A}) = \left|\begin{array}{ccc|cc} 2 & 4 & 1 & 2 & 4 \\ 1 & 5 & 2 & 1 & 5 \\ 3 & 0 & 0 & 3 & 0 \end{array}\right. = 2\cdot 5\cdot 0 + 4\cdot 2\cdot 3 + 1\cdot 1\cdot 0 - 1\cdot 5\cdot 3 - 2\cdot 2\cdot 0 - 4\cdot 1\cdot 0 = 9$$

The determinant of the matrix $\mathbf{A}_{n\times n}$ can be expressed by the folowing so called **Laplace expansions** or **cofactor expansions**.

$$\det(\mathbf{A}) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \qquad \text{expansion along the } i\text{th row}$$
$$\det(\mathbf{A}) = a_{1j}C_{1j} + a_{nj}C_{2j} + \cdots + a_{nj}C_{nj} \qquad \text{expansion along the } j\text{th column}$$

Example 19

Compute the determinant of the matrix a) $\mathbf{A} = \begin{pmatrix} 2 & 4 & 1 \\ 1 & 5 & 2 \\ 3 & 0 & 0 \end{pmatrix}$ b) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 6 & 2 \\ 3 & 4 & 0 \end{pmatrix}$

a) Expansion along the second row:

$$\det(\mathbf{A}) = 1\cdot(-1)^3\cdot\begin{vmatrix} 4 & 1 \\ 0 & 0 \end{vmatrix} + 5\cdot(-1)^4\cdot\begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} + 2\cdot(-1)^5\cdot\begin{vmatrix} 2 & 4 \\ 3 & 0 \end{vmatrix} = -0 + 5\cdot(-3) - 2\cdot(-12) = 9$$

b) Expansion along the third column:

$$\det(\mathbf{A}) = 1\cdot(-1)^4\cdot\begin{vmatrix} 5 & 6 \\ 3 & 4 \end{vmatrix} + 2\cdot(-1)^5\cdot\begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} + 0\cdot(-1)^6\cdot\begin{vmatrix} 1 & 1 \\ 5 & 6 \end{vmatrix} = 1\cdot 2 - 2\cdot 1 + 0\cdot 1 = 0$$

Matrix $\mathbf{A}$ is singular see page 16.

A matrix $\mathbf{A}$ is singular if and only if $\det(\mathbf{A}) = 0$.

Example 20

Compute the determinant of the matrix $\mathbf{A} = \begin{pmatrix} 1 & -2 & 1 & 3 \\ 4 & 0 & 1 & 0 \\ -1 & 0 & 2 & 3 \\ 1 & -2 & 1 & 1 \end{pmatrix}$.

To minimize the effort expand along a row (or a column) which contains the most zeros.
Expansion along the second column:

$$\det(\mathbf{A}) = (-2)\cdot(-1)^3\cdot\begin{vmatrix} 4 & 1 & 0 \\ -1 & 2 & 3 \\ 1 & 1 & 1 \end{vmatrix} + 0 + 0 + (-2)\cdot(-1)^6\cdot\begin{vmatrix} 1 & 1 & 3 \\ 4 & 1 & 0 \\ -1 & 2 & 3 \end{vmatrix} = 2\cdot 0 - 2\cdot 18 = -36$$

Exercise 21

Compute the determinant of the following matrices.

a) $\mathbf{A} = \begin{pmatrix} 5 & 0 & 1 \\ -1 & 3 & 4 \\ 1 & 0 & 2 \end{pmatrix}$ b) $\mathbf{B} = \begin{pmatrix} 4 & 1 & 1 & 0 \\ 3 & 1 & -2 & 1 \\ 0 & 0 & 2 & 1 \\ 2 & 0 & 3 & 1 \end{pmatrix}$ c) $\mathbf{C} = \begin{pmatrix} 0 & 1 & 0 & 5 \\ -3 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & -2 & 0 & 3 \end{pmatrix}$

3 Vector spaces $\mathbb{R}^n$

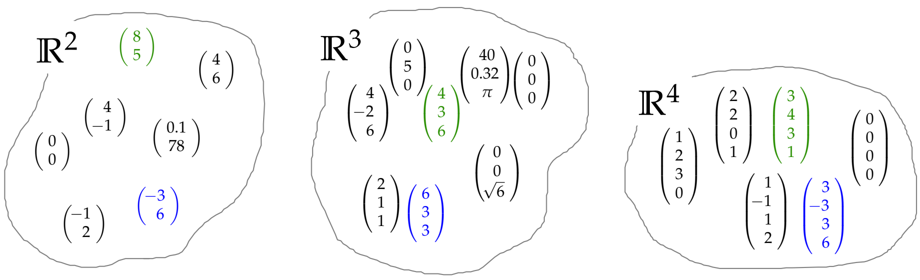
Definition

The **vector space** $\mathbb{R}^n$ consist of all **vectors** $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ with n real entries.

Any two vectors can be added together and any vector can be multiplied by a real number.

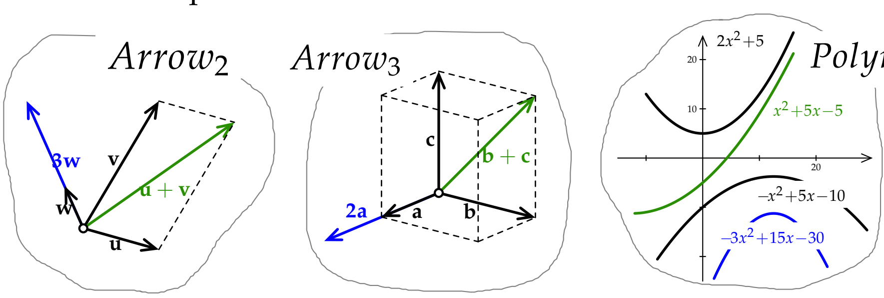
$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix} \qquad c \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} c \cdot u_1 \\ c \cdot u_2 \\ \vdots \\ c \cdot u_n \end{pmatrix}$$

The vector $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ from $\mathbb{R}^n$ consisting of n zeros is called the **zero vector**. It is denoted **o**.



$\mathbb{R}^5$, $\mathbb{R}^6$, ...

Although we will only study the vector spaces $\mathbb{R}^n$, there are other sets of objects that also form vector spaces.

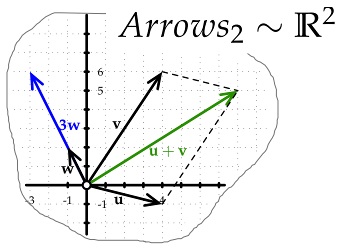


Polynom$_3$, *Polynom*$_4$, ...

and many others...

All their objects (vectors) have some things in common. They can be added together, multiplied by a number, there is a "zero" vector among them, etc.

There is a correspondence between some vector spaces. For example *Arrows*$_2 \sim \mathbb{R}^2$, *Arrows*$_3 \sim \mathbb{R}^3$. There is no corresponding "arrow" space for $\mathbb{R}^4$, $\mathbb{R}^5$, ...



3.1 Linear combination and span

Definition

We say that the vector $\mathbf{w}$ is a **linear combination** of the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$ if there are numbers $c_1, c_2, \ldots c_r$ such that

$$\mathbf{w} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_r\mathbf{v_r}.$$

Definition

For a set of vectors $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ the **span of** S is the set of all linear combinations of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$. It is denoted $\text{span}(S)$.

The vector $\begin{pmatrix}3\\3\\4\end{pmatrix}$ is a linear combination of the vectors $\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\2\\2\end{pmatrix}$ since

$$\begin{pmatrix}3\\3\\4\end{pmatrix} = 1\cdot\begin{pmatrix}0\\0\\1\end{pmatrix} + 1\cdot\begin{pmatrix}1\\1\\1\end{pmatrix} + 1\cdot\begin{pmatrix}2\\2\\2\end{pmatrix} \quad \text{or} \quad \begin{pmatrix}3\\3\\4\end{pmatrix} = 1\cdot\begin{pmatrix}0\\0\\1\end{pmatrix} + 3\cdot\begin{pmatrix}1\\1\\1\end{pmatrix} + 0\cdot\begin{pmatrix}2\\2\\2\end{pmatrix}$$

This can be writtten as a matrix equation:

$$\begin{pmatrix}0&1&2\\0&1&2\\1&1&2\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}3\\3\\4\end{pmatrix} \quad \text{or} \quad \begin{pmatrix}0&1&2\\0&1&2\\1&1&2\end{pmatrix}\begin{pmatrix}1\\3\\0\end{pmatrix} = \begin{pmatrix}3\\3\\4\end{pmatrix}$$

Example 22

Decide whether the vector $\mathbf{w}$ is a linear combination of the vectors $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$;

$$\mathbf{w} = \begin{pmatrix}0\\3\\0\\0\end{pmatrix}, \quad \mathbf{v_1} = \begin{pmatrix}2\\0\\0\\0\end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix}1\\0\\2\\1\end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix}0\\0\\5\\1\end{pmatrix}.$$

We are looking for three numbers c_1, c_2, c_3 such that

$$c_1\cdot\begin{pmatrix}2\\0\\0\\0\end{pmatrix} + c_2\cdot\begin{pmatrix}1\\0\\2\\1\end{pmatrix} + c_3\cdot\begin{pmatrix}0\\0\\5\\1\end{pmatrix} = \begin{pmatrix}0\\3\\0\\0\end{pmatrix}$$

This can be written as the matrix equation below and then solved by elimination.

$$\begin{pmatrix}2&1&0\\0&0&0\\0&2&5\\0&1&1\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} = \begin{pmatrix}0\\3\\0\\0\end{pmatrix} \qquad \left(\begin{array}{ccc|c}\textcircled{2}&1&0&0\\0&0&0&3\\0&2&5&0\\0&1&1&0\end{array}\right) \xrightarrow{R_2\leftrightarrow R_4} \left(\begin{array}{ccc|c}\textcircled{2}&1&0&0\\0&1&1&0\\0&2&5&0\\0&0&0&3\end{array}\right)$$

There is no solution, therefore the vector $\mathbf{w}$ is not a linear combination of the vectors $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$. In other words, the vector $\mathbf{w}$ is not in the span of $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$.

Example 23

Decide whether the vector $\mathbf{w}$ is a linear combination of the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$;

$$\mathbf{w} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 6 \end{pmatrix}, \quad \mathbf{v_1} = \begin{pmatrix} 2 \\ 0 \\ 4 \\ 2 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 2 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$

We are looking for three numbers c_1, c_2, c_3 such that

$$c_1 \cdot \begin{pmatrix} 2 \\ 0 \\ 4 \\ 2 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} + c_3 \cdot \begin{pmatrix} 2 \\ 0 \\ 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 6 \end{pmatrix}$$

This can be written as the matrix equation

$$\begin{pmatrix} 2 & 1 & 2 \\ 0 & 1 & 0 \\ 4 & 2 & 5 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 6 \end{pmatrix}$$

and solved by elimination.

$$\left(\begin{array}{ccc|c} \textcircled{2} & 1 & 2 & 2 \\ 0 & 1 & 0 & 0 \\ 4 & 2 & 5 & 0 \\ 2 & 1 & 1 & 6 \end{array}\right)_{\substack{R_3 - 2R_1 \\ R_4 - R_1}} \rightarrow \left(\begin{array}{ccc|c} \textcircled{2} & 1 & 2 & 2 \\ 0 & \textcircled{1} & 0 & 0 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & -1 & 4 \end{array}\right)_{R_4 + R_3} \rightarrow \left(\begin{array}{ccc|c} \textcircled{2} & 1 & 2 & 2 \\ 0 & \textcircled{1} & 0 & 0 \\ 0 & 0 & \textcircled{1} & -4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Back substitution gives the solution $c_1 = 5, c_2 = 0, c_3 = -4$. Therefore the vector $\mathbf{w}$ is a linear combination of the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$. In other words, the vector $\mathbf{w}$ is in the span of $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$.

$$\mathbf{w} = 5\mathbf{v_1} + 0\mathbf{v_2} - 4\mathbf{v_3}.$$

Exercise 24

a) Decide whether the vector $\mathbf{w}$ is a linear combination of the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$;

$$\mathbf{w} = \begin{pmatrix} 3 \\ 4 \\ 3 \\ -4 \end{pmatrix}, \quad \mathbf{v_1} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 0 \\ 2 \\ -4 \\ -2 \end{pmatrix}.$$

b) Decide whether the vector $\mathbf{w}$ is in the span of $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$;

$$\mathbf{w} = \begin{pmatrix} 5 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v_1} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 0 \\ 2 \\ -4 \\ -2 \end{pmatrix}.$$

3.2 Linear independence

Definition

The sequence of vectors $\langle \mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r} \rangle$ is called **linearly independent** if the only solution to the equation

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_r\mathbf{v_r} = \mathbf{o}$$

is $c_1 = c_2 = \cdots = c_r = 0$.

If there is a solution with at least one nonzero c_i, the sequence of vectors is called **linearly dependent**.

Notice that if vector $\mathbf{v_4}$ is a linear combination of vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$: $\mathbf{v_4} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3}$

then vectors $\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4} \rangle$ are lineary dependent since $c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3} - \mathbf{v_4} = \mathbf{o}$.

Example 25

Decide whether the sequence of vectors $\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4} \rangle$ is linearly independent.

$$\mathbf{v_1} = \begin{pmatrix} 2 \\ 4 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 5 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 3 \\ 4 \\ 3 \\ 7 \end{pmatrix}, \quad \mathbf{v_4} = \begin{pmatrix} 4 \\ 4 \\ 6 \\ 12 \end{pmatrix}.$$

$$\left(\begin{array}{cccc|c} \textcircled{2} & 1 & 3 & 4 & 0 \\ 4 & 0 & 4 & 4 & 0 \\ 0 & 3 & 3 & 6 & 0 \\ 2 & 5 & 7 & 12 & 0 \end{array}\right) \overset{R_2-2R_1}{\underset{R_4-R_1}{\longrightarrow}} \left(\begin{array}{cccc|c} \textcircled{2} & 1 & 1 & 3 & 0 \\ 0 & \textcircled{-2} & -2 & -4 & 0 \\ 0 & 3 & 3 & 6 & 0 \\ 0 & 4 & 4 & 8 & 0 \end{array}\right) \overset{2R_1+R_2}{\underset{\substack{2R_3+3R_2 \\ R_4+2R_2}}{\longrightarrow}} \left(\begin{array}{cccc|c} \textcircled{4} & 0 & 4 & 4 & 0 \\ 0 & \textcircled{-2} & -2 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \overset{R_1/4}{\underset{R_2/-2}{\longrightarrow}} \left(\begin{array}{cccc|c} \textcircled{1} & 0 & 1 & 1 & 0 \\ 0 & \textcircled{1} & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The system has infinitely many solutions $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = t \cdot \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} + s \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$

Therefore the sequence $\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4} \rangle$ is lineary dependent.

Forexample: for $t = 1, s = 0$: $\quad -\mathbf{v_1} - 2\mathbf{v_2} + 0\mathbf{v_3} + \mathbf{v_4} = \mathbf{o}$,

for $t = 1, s = 1$: $\quad -2\mathbf{v_1} - 3\mathbf{v_2} + \mathbf{v_3} + 0\mathbf{v_4} = \mathbf{o}, \quad \ldots$

Example 26

Decide whether the sequence of vectors $\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4} \rangle$ is linearly independent.

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v_4} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$\left(\begin{array}{cccc|c} \textcircled{1} & 0 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{array}\right) \underset{\substack{R_3-R_1 \\ R_4-R_1}}{\longrightarrow} \left(\begin{array}{cccc|c} \textcircled{1} & 0 & 1 & 1 & 0 \\ 0 & \textcircled{3} & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right) \underset{\substack{R_3/-1 \\ R_4/-1}}{\longrightarrow} \left(\begin{array}{cccc|c} \textcircled{1} & 0 & 1 & 1 & 0 \\ 0 & \textcircled{3} & 1 & 1 & 0 \\ 0 & 0 & \textcircled{1} & 0 & 0 \\ 0 & 0 & 0 & \textcircled{1} & 0 \end{array}\right)$$

The system has only the trivial solution: $0\mathbf{v_1} + 0\mathbf{v_2} + 0\mathbf{v_3} + 0\mathbf{v_4} = \mathbf{o}$.

Therefore the sequence $\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4} \rangle$ is lineary independent.

3.3 Metric structure of $\mathbb{R}^n$

Definition

For two vectors $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ of $\mathbb{R}^n$ their **dot product** (or the **standard inner product**) is defined to be the number

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

The dot product of the vectors $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 0 \\ -5 \\ 1 \end{pmatrix}$ is

$$\mathbf{u} \cdot \mathbf{v} = 1 \cdot 0 + 2 \cdot (-5) + 3 \cdot 1 = -7$$

Definition

For the vector $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ of $\mathbb{R}^n$ the **magnitude of vector** (or the **euclidean vector norm**) is defined to be the number $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$

The magnitude of the vector $\mathbf{u} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 4 \end{pmatrix}$ is $\|\mathbf{u}\| = \sqrt{3^2 + 1^2 + 4^2} = \sqrt{26} \doteq 5.09$

Definition

For two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ of $\mathbb{R}^n$ the **angle** between them is defined to be the number $\varphi \in \langle 0, \pi \rangle$ such that

$$\cos \varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

Find the angle of the vectors $\mathbf{u} = \begin{pmatrix} 5 \\ -4 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$.

$$\cos \varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{5 \cdot 2 + (-4) \cdot 1 + 3 \cdot 3}{\sqrt{5^2 + (-4)^2 + 3^2}\sqrt{2^2 + 1^2 + 3^2}} = \frac{15}{\sqrt{50}\sqrt{14}} \doteq 0.5669$$

$$\varphi = 0.96 \text{ rad} \qquad \varphi = 55.46^\circ$$

Definition

Two vectors of $\mathbb{R}^n$ are said to be **orthogonal** or **perpendicular** (to each other) if their dot product equals zero.

$$\mathbf{u} \cdot \mathbf{v} = 0$$

Exercise 27

a) Find the dot product of the vectors $\mathbf{u} = \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix}$.

b) Find the magnitude of the vector $\mathbf{u} = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}$.

c) Find the angle of the vectors $\mathbf{u}$ and $\mathbf{v}$.

i) $\mathbf{u} = \begin{pmatrix} 4 \\ -1 \\ 5 \\ 2 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} 1 \\ 3 \\ -2 \\ 2 \end{pmatrix}$

ii) $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} -2 \\ 3 \\ -2 \end{pmatrix}$

d) Decide, whether the vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

i) $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$

ii) $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} -2 \\ 3 \\ -2 \end{pmatrix}$

e) Fill in the missing numbers, so that the vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

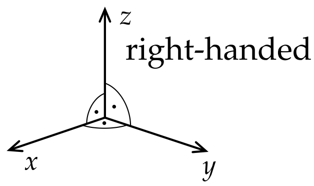
i) $\mathbf{u} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} -5 \\ 7 \\ * \end{pmatrix}$

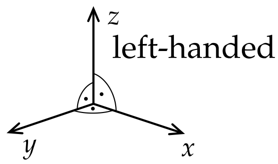
ii) $\mathbf{u} = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 4 \end{pmatrix}$ $\mathbf{v} = \begin{pmatrix} * \\ 4 \\ 2 \\ 1 \end{pmatrix}$

4 Euclidean space $\mathbb{E}_3$

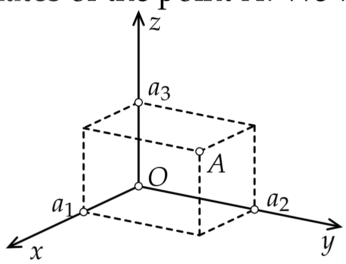
Definition

A Cartesian coordinate system (in 3D) consists of an ordered triplet of orientated lines (**the axes**) pair-wise perpendicular that go through a common point (**the origin**) and are pair-wise perpendicular; and a single unit of length common for all three axes.
The axes are denoted x, y, z. The xy-plane, yz-plane, xz-plane are called **coordinate planes**. The system can be either **right-handed** or **left-handed**.





In 3D space equipped with a Cartesian coordinate system, every point A is uniquely determined by an ordered triplet of numbers $[a_1, a_2, a_3]$ as shown in the picture below. The numbers are called **coordinates** of the point A. We write this as $A = [a_1, a_2, a_3]$.



From now on to conserve space we will write vectors from the space $\mathbb{R}^3$ differently. Instead of $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ we will write $\mathbf{u} = (u_1, u_2, u_3)$ (numbers "laying down" and divided by commas).

Definition

Euclidean space $\mathbb{E}_3$ contains two sets of object. The set of all points $[a_1, a_2, a_3]$ and the set of all vectors (u_1, u_2, u_3) from the vector space $\mathbb{R}^3$ equipped with the dot product. Both sets are "tied up" together, we can "add" point to a vector to get another point

$$A + \mathbf{u} = [a_1 + u_1, a_2 + u_2, a_3 + u_3].$$

For every (ordered) pair of points $A = [a_1, a_2, a_3]$ and $B = [b_1, b_2, b_3]$ there is a unique vector

$$\mathbf{AB} = (b_1 - a_1, b_2 - a_2, b_3 - a_3).$$

The distance of two points A and B is the magnitude of the vector $\mathbf{AB}$.

Example 28

For the points $A = [2, -1, 5]$, $B = [4, 1, 0]$ find the vector $\mathbf{AB}$.

$$\mathbf{AB} = (4 - 2, 1 - (-1), 0 - 5) = (2, 2, -5)$$

Recall that two vectors are perpendicular whenever their dot product is zero.
Vectors perpendicular to the vector $\mathbf{u}=(1,0,2)$ are for example $(-2,0,1)$, $(4,5,-2)$ ...,
vectors perpendicular to the vector $\mathbf{v}=(3,4,1)$ are for example $(1,-1,1)$, $(0,1,-4)$ ....
Is there a way to find a vector that will be perpendicular to both of them?

Definition

For two vectors $\mathbf{u}=(u_1,u_2,u_3)$ and $\mathbf{v}=(v_1,v_2,v_3)$ of $\mathbb{R}^3$ their **cross product** is defined to be the vector

$$\mathbf{u}\times\mathbf{v}=(u_2v_3-u_3v_2,\ u_3v_1-u_1v_3,\ u_1v_2-u_2v_1)$$

The cross product $\mathbf{u}\times\mathbf{v}$ is perpendicular to the vector $\mathbf{u}$ and to the vector $\mathbf{v}$.

The cross product can also be expressed as the formal determinant

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix}=\mathbf{i}\cdot\begin{vmatrix}u_2 & u_3\\ v_2 & v_3\end{vmatrix}-\mathbf{j}\cdot\begin{vmatrix}u_1 & u_3\\ v_1 & v_3\end{vmatrix}+\mathbf{k}\cdot\begin{vmatrix}u_1 & u_2\\ v_1 & v_2\end{vmatrix}$$

where $\mathbf{i}=(1,0,0)$, $\mathbf{j}=(0,1,0)$, $\mathbf{k}=(0,0,1)$.

Example 29

Find the cross product of the vectors $\mathbf{u}=(1,0,2)$ and $\mathbf{v}=(3,4,1)$. Check that the cross product is perpendicular to the vector $\mathbf{u}$ and to the vector $\mathbf{v}$.

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & 0 & 2\\ 3 & 4 & 1\end{vmatrix}=\left(\begin{vmatrix}0 & 2\\ 4 & 1\end{vmatrix},\ -\begin{vmatrix}1 & 2\\ 3 & 1\end{vmatrix},\ \begin{vmatrix}1 & 0\\ 3 & 4\end{vmatrix}\right)=(-8,5,4)$$

To check whether they are perpendicular we calculate their dot product:
$(1,0,2)\cdot(-8,5,4)=-8+0+8=0$ Vectors $\mathbf{u}$ and $\mathbf{u}\times\mathbf{v}$ are perpendicular.
$(3,4,1)\cdot(-8,5,4)=-24+20+4=0$ Vectors $\mathbf{v}$ and $\mathbf{u}\times\mathbf{v}$ are perpendicular.

Exercise 30

a) Find the cross product of the vectors $\mathbf{u}=(4,-1,2)$ and $\mathbf{v}=(2,0,5)$. Check that the cross product is perpendicular to the vector $\mathbf{u}$ and to the vector $\mathbf{v}$.

b) For the point $A=[3,4,1]$ and the vector $\mathbf{u}=(2,-2,1)$ find the coordinates of the following points:

i) $A+\mathbf{u}$, ii) $A+2\mathbf{u}$, iii) $A+3\mathbf{u}$, iv) $A-\mathbf{u}$.

4.1 Line in $\mathbb{E}_3$

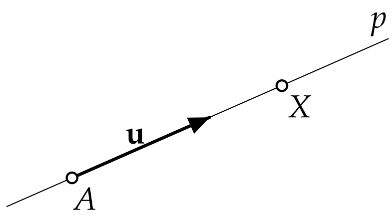
Definition

A **line** through a point $A=[a_1,a_2,a_3]$ in the direction of a vector $\mathbf{u}=(u_1,u_2,u_3)$ is defined to be a set of all points X satisfying equation

$$X = A + t\mathbf{u}, \quad t \in \mathbb{R}$$

The equations

$$\begin{aligned} x &= a_1 + tu_1 \\ p: y &= a_2 + tu_2 \\ z &= a_3 + tu_3, \quad t \in \mathbb{R}. \end{aligned}$$



are called the **parametric equations of the line**.

For example the parametric equations of the line passing through the point $A=[3,4,1]$ in the direction of the vector $\mathbf{u}=(2,-2,1)$ are:

$$\begin{aligned} x &= 3 + 2t \\ p: y &= 4 - 2t \\ z &= 1 + t, \quad t \in \mathbb{R} \end{aligned}$$

Exercise 31

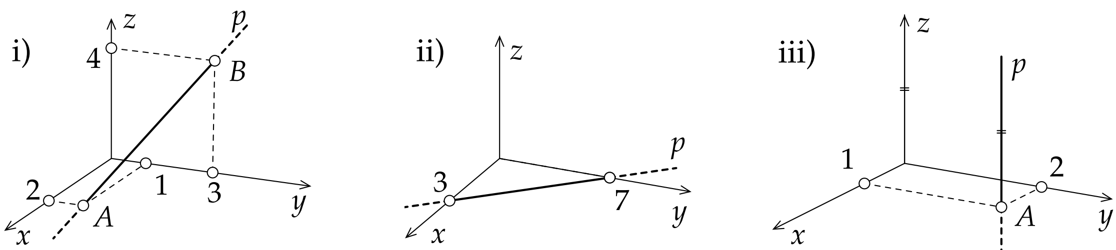
a) Find the parametric equations for a line p which passes through the point $A=[1,4,-1]$ in the direction of the vector $\mathbf{u}=(3,0,2)$.

   i) Find the coordinates of any four points the line p passes through.

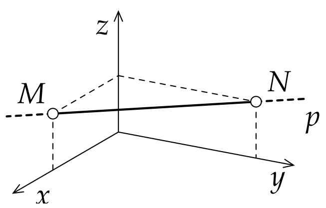
   ii) Find the missing coordinates of points $[-2, *, *]$, $[*, *, 7]$ and $[*, 6, *]$ so they lie on the line p.

   iii) Find out whether the line p passes through the point $A[4,4,4]$, $B[10,4,5]$.

b) Find the parametric equations for the line p in the pictures below.



c) Find the coordinates of the point M and the point N lying on the line p and also in the xz-plane and yz-plane, respectively.



$$\begin{aligned} x &= 8 - 2t \\ p: y &= -9 + 3t \\ z &= 5 \end{aligned}$$

d) Decide whether there is a line which passes through all three points $A=[1,1,1]$, $B=[4,3,5]$ and $C=[7,5,9]$. If so, write its parametric equations.

4.2 Line – line intersection

Definition

If there is a point lying on two lines it is called their **point of intersection**.

Example 32

Find the point of intersection of the lines p and q (if there is any).

$$p: \begin{aligned} x &= 8-3t \\ y &= 1-t \\ z &= 3+2t \end{aligned} \qquad q: \begin{aligned} x &= 3+s \\ y &= s \\ z &= 8+s \end{aligned}$$

Let's assume there is a point of intersection $P = [p_1, p_2, p_3]$. Since P lies on both lines, its coordinates must satisfy both equations for some t and some s.

$$\begin{aligned} p_1 &= 8-3t \\ p_2 &= 1-t \\ p_3 &= 3+2t \end{aligned} \qquad \begin{aligned} p_1 &= 3+s \\ p_2 &= s \\ p_3 &= 8+s \end{aligned} \qquad \begin{aligned} 8-3t &= 3+s \\ 1-t &= s \\ 3+2t &= 8+s \end{aligned}$$

We get three linear equations in two unknowns.

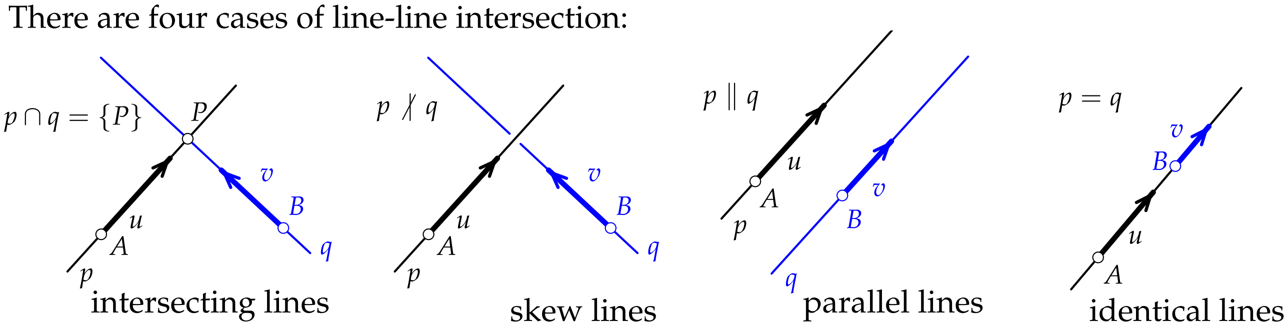
$$\begin{aligned} -3t-s &= -5 \\ -t-s &= -1 \\ 2t-s &= 5 \end{aligned} \qquad \left(\begin{array}{rr|r} -3 & -1 & -5 \\ 0 & 2 & -2 \\ 0 & -5 & 5 \end{array}\right) \to \left(\begin{array}{rr|r} -3 & -1 & -5 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{array}\right)$$

The system has one solution $t = 2$ a $s = -1$. The lines intersect at one point.

$$\begin{aligned} p_1 &= 8-3t = 8-3\cdot 2 = 2 \\ p_2 &= 1-t = 1-2 = -1 \\ p_3 &= 3+2t = 3+2\cdot 2 = 7 \end{aligned} \qquad \begin{aligned} p_1 &= 3+s = 3+(-1) = 2 \\ p_2 &= s = -1 \\ p_3 &= 8+s = 8-1 = 7 \end{aligned}$$

Their point of intersection is $P = [2, -1, 7]$.

There are four cases of line-line intersection:



Example 33

Decide whether the lines p, q are identical, parallel, skewed or intersecting.

a) $p: \begin{array}{l} x = 2 + t \\ y = 3 + t \\ z = 5 + 2t \end{array}$ $\quad q: \begin{array}{l} x = -3 + r \\ y = 6 - r \\ z = 7 - r \end{array}$ $\qquad$ c) $p: \begin{array}{l} x = 1 + t \\ y = 1 - 2t \\ z = -2 + 3t \end{array}$ $\quad q: \begin{array}{l} x = 2 + 2r \\ y = 2 - 4r \\ z = -5 + 6r \end{array}$

b) $p: \begin{array}{l} x = 1 + t \\ y = 1 - 2t \\ z = -2 + 3t \end{array}$ $\quad q: \begin{array}{l} x = 2 - r \\ y = 2 + 3r \\ z = -5 - 4r \end{array}$ $\qquad$ d) $p: \begin{array}{l} x = 1 + t \\ y = 1 - 2t \\ z = -2 + 3t \end{array}$ $\quad q: \begin{array}{l} x = 2 - 3r \\ y = -1 + 6r \\ z = 1 - 9r \end{array}$

a)
$$\begin{array}{rl} 2 + t &= -3 + r \\ 3 + t &= 6 - r \\ 5 + 2t &= 7 - r \end{array} \qquad \begin{array}{rl} t - r &= -5 \\ t + r &= 3 \\ 2t + r &= 2 \end{array} \qquad \left(\begin{array}{rr|r} 1 & -1 & -5 \\ 1 & 1 & 3 \\ 2 & 1 & 2 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & -1 & -5 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{array}\right)$$

The system has one solution $r = 4$, $t = -1$. The **lines intersect** in one point $[1, 2, 3]$.

b)
$$\begin{array}{rl} 1 + t &= 2 - r \\ 1 - 2t &= 2 + 3r \\ -2 + 3t &= -5 - 4r \end{array} \qquad \begin{array}{rl} t + r &= 1 \\ -2t - 3r &= 1 \\ 3t + 4r &= -3 \end{array} \qquad \left(\begin{array}{rr|r} 1 & 1 & 1 \\ -2 & -3 & 1 \\ 3 & 4 & -3 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 1 & 1 \\ 0 & -1 & 3 \\ 0 & 0 & -9 \end{array}\right)$$

The system has no solution. The lines don't intersect. Thus they are either parallel or skew. Since their direction vectors $(1, -2, 3)$ and $(-1, 3, -4)$ don't have same direction, the **lines are skewed**.

c)
$$\begin{array}{rl} 1 + t &= 2 + 2r \\ 1 - 2t &= 2 - 4r \\ -2 + 3t &= -5 + 6r \end{array} \qquad \begin{array}{rl} t - 2r &= 1 \\ -2t + 4r &= 1 \\ 3t - 6r &= -3 \end{array} \qquad \left(\begin{array}{rr|r} 1 & -2 & 1 \\ -2 & 4 & 1 \\ 3 & -6 & -3 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & -2 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{array}\right)$$

The system has no solution. The lines don't intersect. Therefore they are either parallel or skewed. Since their direction vectors $(1, -2, 3)$ and $(2, -4, 6)$ have the same direction, the **lines are parallel**.

d)
$$\begin{array}{rl} 1 + t &= 2 - 3r \\ 1 - 2t &= -1 + 6r \\ -2 + 3t &= 1 - 9r \end{array} \qquad \begin{array}{rl} t + 3r &= 1 \\ -2t - 6r &= -2 \\ 3t + 9r &= 3 \end{array} \qquad \left(\begin{array}{rr|r} 1 & 3 & 1 \\ -2 & -6 & -2 \\ 3 & 9 & 3 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 3 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

The system has infitely many solutions. The lines have infinitely many common points. Therefore they are **identical**.

Exercise 34

Decide whether the lines p, q are identical, parallel, skewed or intersecting.

a) $p: \begin{array}{l} x = 1 + 2t \\ y = 7 - 6t \\ z = -2 + 8t \end{array}$ $\quad q: \begin{array}{l} x = -2 - s \\ y = 10 + 3s \\ z = 1 - 4s \end{array}$ $\qquad$ c) $p: \begin{array}{l} x = 1 + t \\ y = 1 - 2t \\ z = 2 + 2t \end{array}$ $\quad q: \begin{array}{l} x = 2 - s \\ y = -1 + s \\ z = 1 - s \end{array}$

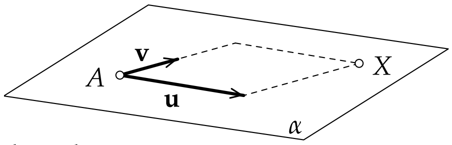
b) $p: \begin{array}{l} x = 7 + t \\ y = -11 + 3t \\ z = -10 + 3t \end{array}$ $\quad q: \begin{array}{l} x = 5 + 6s \\ y = 3 - 2s \\ z = 10 + s \end{array}$ $\qquad$ d) $p: \begin{array}{l} x = 10 + t \\ y = -7 - t \\ z = 3 + 3t \end{array}$ $\quad q: \begin{array}{l} x = 14 - 2s \\ y = 3 + 2s \\ z = 15 - 6s \end{array}$

4.3 Plane in $\mathbb{E}_3$

Definition

A **plane** passing through the point $A = [a_1, a_2, a_3]$ in the direction of two independent vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ is defined to be the set of all points X such that

$$X = A + t\mathbf{u} + s\mathbf{v}$$



for some real numbers t, s.

The vectors **u** and **v** are called the **direction vectors** of the plane α.

The equations

$$\begin{aligned} x &= a_1 + tu_1 + sv_1 \\ \alpha : y &= a_2 + tu_2 + sv_2 \\ z &= a_3 + tu_3 + sv_3 \qquad t, s \in \mathbb{R}. \end{aligned}$$

are called the **parametric equations of the plane**.
A vector is said to **lie in the plane** if it is a linear combination of **u** and **v**.

For the example parametric equations of the plane α which passes through the point $A = [1, 2, -2]$ in the direction of $\mathbf{u} = (3, -1, -1)$ and $\mathbf{v} = (1, 0, -2)$ are

$$\begin{aligned} x &= 1 + 3t + s \\ y &= 2 - t \\ z &= -2 - t - 2s \qquad t, s \in \mathbb{R}. \end{aligned}$$

To eliminate the parametrs t and s we multiply first equation by two, the second by five and add all of the equations together.

$$\begin{aligned} 2x &= 2 + 6t + 2s \\ 5y &= 10 - 5t \\ z &= -2 - t - 2s \\ \hline 2x + 5y + z &= 10 + 0t + 0s \end{aligned}$$

The result is $\quad 2x + 5y + z - 10 = 0$

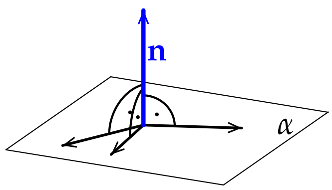
Definition

The equation $\qquad \alpha : \quad ax + by + cz + d = 0$

is called the **general equation** of the plane. The numbers a, b, c must not all be zero.

Definition

A vector **n** is said to be **perpendicular to a plane** if it is perpendicular to all vectors that lie in the plane. Any such vector is called a **normal vector** of the plane.



A normal vector of the plane $\alpha : \quad ax + by + cz + d = 0$ is the vector (a, b, c).

A normal vector of the plane $\alpha : 2x + 5y + 1z - 10 = 0$ is the vector $\mathbf{n} = (2, 5, 1)$.
Notice that the vector **n** is perpendicular to vector **u** and vector **v**.

Example 35

Check whether the point $A = [1, 2, -2]$ lies in the plane $\alpha: 2x + 5y + z - 10 = 0$.

$2 \cdot 1 + 5 \cdot 2 + 1 \cdot (-2) - 10 = 0$. Yes, it does.

Example 36

Find a general equation of the plane that goes through the point $A = [1, 2, -2]$ in direction of vectors $\mathbf{u} = (3, -1, -1)$ and $\mathbf{v} = (1, 0, -2)$.

The equation we are looking for is in the form $ax + by + cz + d = 0$, where (a, b, c) is a normal vector. A normal vector is one that is perpendicular to $\mathbf{u}$ and $\mathbf{v}$, so their vector product $\mathbf{u} \times \mathbf{v}$ is a normal vector.

$$(a, b, c) = \mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} -1 & -1 \\ 0 & -2 \end{vmatrix}, - \begin{vmatrix} 3 & -1 \\ 1 & -2 \end{vmatrix}, \begin{vmatrix} 3 & -1 \\ 1 & 0 \end{vmatrix} \right) = (2, 5, 1)$$

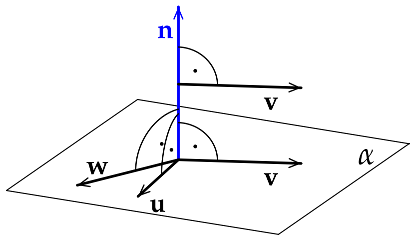
We know that $a = 2$, $b = 5$ and $c = 1$. To find the number d we use the fact that the point A lies in the plane. Therefore

$$\begin{aligned} 2 \cdot (1) + 5 \cdot (2) + (-2) + d &= 0 \\ d &= -10 \end{aligned}$$

A general equation of the plane is $\quad 2x + 5y + z - 10 = 0$.

Example 37

Do the vectors $(1, 5, 9)$, $(6, 4, 3)$ lie in the plane $\alpha: x - 3y + 2z + 2 = 0$?
Write a few other vectors which lie in the plane.



Every vector which lies in a plane is perpendicular to every normal vector of the plane. This can be tested by their dot product.

$(1, 5, 9) \cdot (1, -3, 2) \neq 0$ — the vector $(1, 5, 9)$ is not in the plane α

$(6, 4, 3) \cdot (1, -3, 2) = 0$ — the vector $(6, 4, 3)$ lies in the plane α

The vectors $(7, 1, -2)$, $(9, 5, 3)$, $(1, 1, 1)$, $(2, 0, -1)$ also lie in the plane α.

Exercise 38

a) The points $[*,\ 0,\ 0]$, $[0,\ 0,\ *]$ lie in the plane $\alpha : 2x + 5y + z - 10 = 0$. Find their missing coordinates.

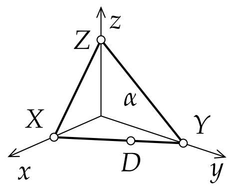
b) Find a general equation of the plane with a normal vector $(-3, 1, 4)$ and passing through point $[1, 2, 1]$.

c) Find a general equation of the plane that goes through the point $A = [-2, 0, 5]$ in the direction of the vectors $\mathbf{u} = (4, 2, -1)$ a $\mathbf{v} = (-1, 1, 2)$. Check whether points $D = [5, 5, 5]$ and $E = [6, 6, 6]$ lie in the plane.

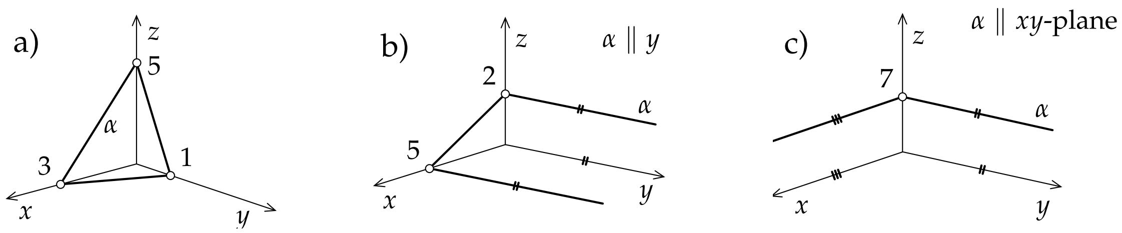
d) Find a general equation of the plane that goes through the points $A = [3, 1, 5]$, $B = [4, 2, 7]$ and $C = [5, 3, 9]$. Find the coordinates of any four points lying in the plane.

e) Find a general equation of the plane α perpendicular to the x-axis and passing through the point $Q = [1, -2, 3]$.

f) Find the coordinates of points X, Y and Z lying on the axes and in the plane $\alpha : 4x + 6y + z - 12 = 0$ and the coordinates of any point D lying in the plane α and the xy-plane.



g) Find a general equation of the plane α in the pictures below.



4.4 Plane – plane intersection

Example 39

Find the line of intersection between the two planes α, β (if there is any).

$$\alpha : x - y + 4z + 2 = 0 \qquad \beta : 2x - y + 5z - 2 = 0$$

Let's assume that $P = [p_1, p_2, p_3]$ is some common point of both planes. Since P lies on both planes, its coordinates must satisfy both equations.

$$\begin{aligned} p_1 - p_2 + 4p_3 + 2 &= 0 \\ 2p_1 - p_2 + 5p_3 - 2 &= 0 \end{aligned}$$

We get a system of two equations with three unknowns.

$$\left(\begin{array}{ccc|c} 1 & -1 & 4 & -2 \\ 2 & -1 & 5 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 4 & -2 \\ 0 & 1 & -3 & 6 \end{array}\right) \qquad \begin{aligned} p_1 &= 4 - t \\ p_2 &= 6 + 3t \\ p_3 &= t \end{aligned}$$

The system has infinitely many solutions. The two planes have infinitely many common points. They all lie on the line p.

$$p : \begin{aligned} x &= 4 - t \\ y &= 6 + 3t \\ z &= t \end{aligned}$$

The line p goes through the point $[4, 6, 0]$ in the direction of the vector $(-1, 3, 1)$. Verify that both the point and the vector lie in both of the planes.

Note: Taking different steps during elimination might get you a different looking solution, for example

$$p : \begin{aligned} x &= 6 - 2s \\ y &= 6s \\ z &= -2 + 2s \end{aligned}$$

These are also equations of the line p.

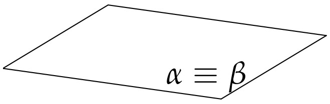
Example 40

Do the points $[1, 1, 1]$, $[1, 0, 4]$, $[2, 1, 0]$ lie in the plane α? Do they lie in the plane β?

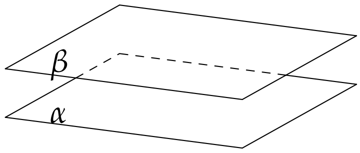
$$\alpha : x + 3y + z - 5 = 0 \qquad \beta : 2x + 6y + 2z - 10 = 0$$

All three points lie in both planes.

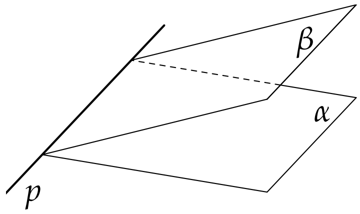
There are three possible cases of plane–plane intersection:



identical planes



parallel planes
contain no common points



intersecting planes

Example 41

Decide whether the planes α and β are identical, intersecting or parallel.

a) $\alpha: 2x+6y+4z+10=0$, $\beta: 3x+9y+6z+15=0$

b) $\alpha: 2x-6y+4z+10=0$, $\beta: 3x-9y+6z+10=0$

c) $\alpha: 2x+6y+4z-10=0$, $\beta: 2x+7y+6z-15=0$

1. method: Compare their equations

a) If the equation of the plane α is a multiple of the equation of the plane β, the planes are **identical**.

b) If the equation of the plane α is a multiple of the equation of the plane β except for the coeficient d, the planes are **parallel**.

c) If the equation of the plane α is not a multiple of the equation of the plane β, the planes are **intersecting**.

2. method: Find common points

a) $$\left(\begin{array}{ccc|c} 2 & 6 & 4 & -10 \\ 3 & 9 & 6 & -15 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 2 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \qquad \alpha \equiv \beta: \begin{array}{l} x = -5 - 2t - 3s \\ y = s \\ z = t \end{array}$$

There are many solutions. Since the system has two free variables, the planes are **identical**. The solution corresponds to the parametric equations of the planes.

b) $$\left(\begin{array}{ccc|c} 2 & -6 & 4 & -10 \\ 3 & -9 & 6 & -10 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -3 & 2 & -5 \\ 0 & 0 & 0 & -25 \end{array}\right)$$

There is no solution, therefore there are no common points. The planes are **parallel**.

c) $$\left(\begin{array}{ccc|c} 2 & 6 & 4 & 10 \\ 2 & 7 & 6 & 15 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 1 & 2 & 5 \end{array}\right) \qquad p: \begin{array}{l} x = -10 + 4t \\ y = 5 - 2t \\ z = t \end{array}$$

There are many solutions. They all lie on a line, the planes are therefore **intersecting**. The solution corresponds to the parametric equations of this line.

Exercise 42

Decide whether the planes α and β are identical, intersecting or parallel.

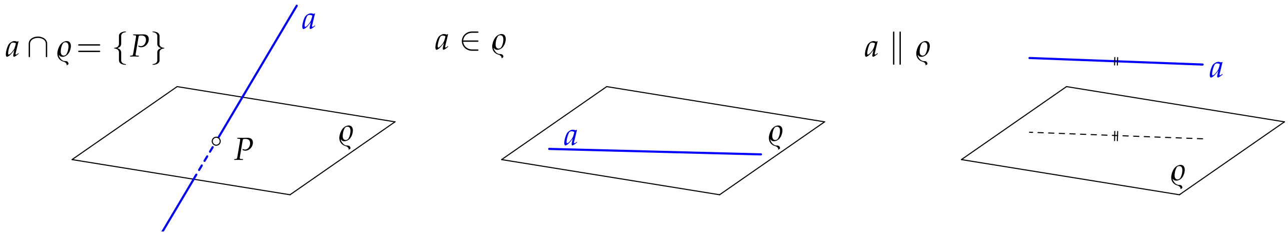
a) $\alpha: x - y + 2z + 2 = 0$, $\beta: 3x-3y+6z+6=0$

b) $\alpha: x - y + 2z + 2 = 0$, $\beta: 5x-5y+10z+3=0$

c) $\alpha: x - y + 2z + 2 = 0$, $\beta: x - 3y + 4z - 4 = 0$

4.5 Plane – line intersection

There are three possible cases of plane-line intersection.



Example 43

Decide whether the given line cuts through, is embeded in or parallel with the plane ϱ.

$$\varrho : 3x - 2y - z - 13 = 0$$

$$a : \begin{aligned} x &= 2 + 2t \\ y &= 1 - t \\ z &= 5t \end{aligned} \qquad b : \begin{aligned} x &= 4 + t \\ y &= 3t \\ z &= -1 - 3t \end{aligned} \qquad c : \begin{aligned} x &= 2 + 2t \\ y &= 1 + 4t \\ z &= 1 - 2t \end{aligned}$$

1. method: Compare the direction vector of the line and normal vector of the plane

a) $(2,-1,5) \cdot (3,-2,-1) \neq 0$ Vector $(2,-1,5)$ does not lie in the plane.
The line cuts through the plane at one single point.

b) $(1,3,-3) \cdot (3,-2,-1) = 0$ Vector $(2,-1,5)$ lies in the plane
Does the point $[4,0,-1]$ of the line lie in the plane? $3 \cdot 4 - 2 \cdot 0 + 1 - 13 = 0$ Yes.
The line is embeded in the plane.

c) $(2,4,-2) \cdot (3,-2,-1) = 0$ Vector $(2,-1,5)$ lies in the plane
Does the point $[2,1,1]$ of the line lie in the plane? $3 \cdot 2 - 2 \cdot 2 - 1 - 13 \neq 0$ No.
The line is parallel with the plane.

2. method: Find common points

a) $$\begin{aligned} p_1 &= 2 + 2t \\ p_2 &= 1 - t \\ p_3 &= 5t \end{aligned} \qquad 3p_1 - 2p_2 - p_3 - 13 = 0 \qquad \begin{aligned} 3(2+2t) - 2(1-t) - 5t - 13 &= 0 \\ t &= 3 \end{aligned}$$

The system has one solution, therefore **the line cuts through the plane at one single point** $P = [8,-2,15]$.

b) $$\begin{aligned} p_1 &= 4 + t \\ p_2 &= 3t \\ p_3 &= -1 - 3t \end{aligned} \qquad 3p_1 - 2p_2 - p_3 - 13 = 0 \qquad \begin{aligned} 3(4+t) - 2(3t) - (-1-3t) - 13 &= 0 \\ 0 &= 0 \end{aligned}$$

The system has many solutions, therefore the line and plane share many common points. **The line is embeded in the plane.**

c) $$\begin{aligned} p_1 &= 2 + 2t \\ p_2 &= 1 + 4t \\ p_3 &= 1 - 2t \end{aligned} \qquad 3p_1 - 2p_2 - p_3 - 13 = 0 \qquad \begin{aligned} 3(2+2t) - 2(1+4t) - (1-2t) - 13 &= 0 \\ -10 &\neq 0 \end{aligned}$$

The system has no solution, therefore the line and plane do not intersect. **The line is parallel with the plane.**

Exercise 44

a) Decide whether the given line cuts through, is embeded in or parallel with the plane ϱ.

$$\varrho : 3x - y + z + 2 = 0$$

$$a : \begin{aligned} x &= 5 - t \\ y &= -1 + 2t \\ z &= -2 + t \end{aligned} \qquad b : \begin{aligned} x &= t \\ y &= 1 + 4t \\ z &= -1 + t \end{aligned} \qquad c : \begin{aligned} x &= 4 + t \\ y &= 2 + t \\ z &= 7 - 2t \end{aligned}$$

b) Decide whether the line a cuts through, is embeded or is parallel with the plane ϱ.

$$a : \begin{aligned} x &= -t - 2 \\ y &= -2t + 4 \\ z &= -2t - 1 \end{aligned} \qquad \varrho : \begin{aligned} x &= -2 + 3s + r \\ y &= 4 + 2s + 2r \\ z &= 1 + 2r \end{aligned}$$

Workbook for Mathematics I

Calculus

Jan Kotůlek

5 Computation of the derivative

5.1 Introduction to Calculus

We begin with a discussion of the two related problems that motivated the invention of calculus. Let us look at the relation between the speedometer and the odometer familiar to every driver. The first one measures velocity v and the other one the distance s travelled.

The relation between v and s.
Can we find v if we know s? How? And vice versa, if we have the record of the velocity over the time, can we compute the distance traveled? In other words, can we recover missing information of odometer form the complete recors of speedometer?

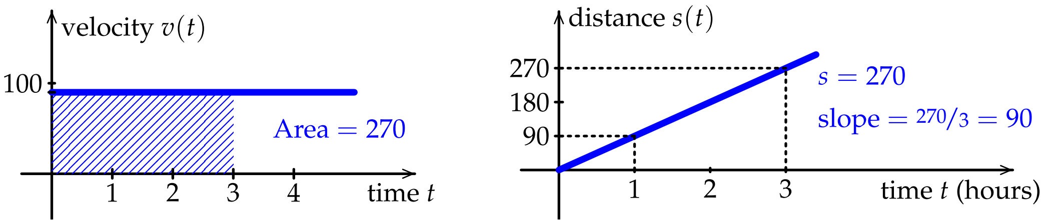
The problem of finding velocity from a record of distance is called **differentiation**, finding distance traveled from the velocity is called **integration**.

Example 45

Constant velocity
Suppose we travel with fixed velocity $v = 90$ (km per hour). Then s incerases at this constant rate. After an hour the distance is $s = 90$ km, after three hours $s = 270$ and after t hours $s = 90t$. The distance incerases **linearly** with time and its graph is a line with slope 90. The graph of velocity is a horizontal line.

This simple relation of v, s, t needs just algebra:

$$s = v \cdot t$$



Constant velocity and linearly increasing distance

Conversely, if s increases linearly, v is constant. The division by time gives the slope. At any point, the ratio s/t is 90. Geometrically, the velocity is **the slope** of the distance graph

$$\text{slope} = \frac{\text{change in distance}}{\text{change in time}} = \frac{v \cdot t}{t} = v$$

Now, we compute s from v. Starting from the graph of v, we discover the graph of $s = v \cdot t$. The graph of s is given by **the area** under the velocity graph. When v is constant, we got a rectangle with height v and width t, hence its area is v times t. Finding area is called **integration**.

- The slope of distance graph gives the velocity v.
- The area under the velocity graph gives the distance s.

Functions

Definition

The **function** is a rule that assigns one member of the range to each member of the domain. Equivalently, we say that a function is a set of ordered pairs $(t, f(t))$ with no t appearing more than once.
The **domain** of the function f is the set D of inputs, $D \subset \mathbb{R}$.
The **range** of the function f is the set I of outputs, $I \subset \mathbb{R}$. We also say that I is **image** of D, $I = f(D)$.

The number $v(t)$, we say "v of t" is the value of the function v at the time t.
The time t is the **input** to the function, the velocity $v(t)$ at that time is the **output**.
The input t is mapped to the output $s(t)$, which changes as t changes. All calculus is about the **rate of change**. This rate was our function $v(t)$.
In some way, functions are instructions telling us how to find s at time t. The instructions can be given in the form of

- **explicit formula** $s = f(t)$, e.g. $s = 2t$,
- **implicit equation** $f(x, y) = 0$, e.g. $x + y - 1 = 0$,
- **parametric equations** $x = x(t)$, $y = y(t)$, with $t \in I \subset \mathbb{R}$, describing coordinates of the point in the plane at the time t, e.g.

$$x(t) = 3 + 3t,$$
$$y(t) = 3 - 3t, \quad t \in \langle 0, 1 \rangle.$$

- **table**

t	0	1	2	3	4	5	6
$s(t)$	0	90	180	270	180	90	0

- **graph**, etc.

In practice, the number $f(t)$ is produced from the number t by reading a graph or display of a measuring device, plugging into a formula, solving an equation, or running a computer program.

There are two central questions leading in opposite directions that calculus was invented to solve:
1) If the velocity is changing, *how can we compute the distance travelled*?
2) If the graph $s(t)$ of the distance is not a straight line, *what is its slope*?

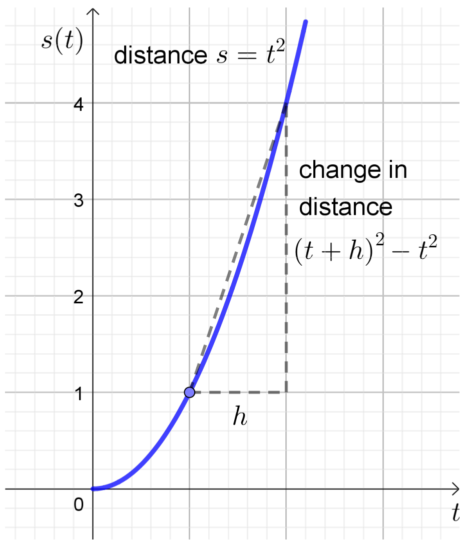
Example 46

Suppose that $s(t) = t^2$ is the distance traveled by time t. Find the velocity $v(t)$.

The distance graph of $s(t)$ is a **parabola**.
Velocity is distance divided by time, but what happens when the speed is changing? Dividing $s = 100$ by $t = 10$ gives $v = 10$, this is the **average velocity** over the first ten seconds. But how do we find the **instantaneous velocity** without looking at the speedometer at the exact instant when $t = 10$?
The problem is that the distance is not distributed evenly. As the car goes faster, the graph of t^2 gets steeper and more distance is covered in every following second. We can try an approximation. The average velocity between $t = 10$ and $t = 11$ may be a good approximation to the speed at the moment $t = 10$ and averages are easy to find:

$$\text{average } v = \frac{\text{change in } s}{\text{change in } t} = \frac{s(11) - s(10))}{11 - 10} = \frac{121 - 100}{1} = 21.$$

The car covered 21 meters in that second and its average speed was 21 m/s. Since it was still gaining speed, the velocity at the beginning $t = 10$ was below 21.
What is the average geometrically? It is a slope. But not the slope of the curve. *The average velocity is the slope of a straight line* joining two point on the curve. Thus, we pretend the velocity is constant and we are back in the previous easy case.



The graph of quadratic distance function $s(t) = t^2$ and its velocity.

We can also find the average over a smaller time interval. The way to find $v(10)$ is to proceed with *reducing the time interval*.

Finding the slope between points that are closer and closer on the curve is the key to the differentail calculus. The "limit" is the slope at a single point.

We can compute the average velocity between $t = 10$ and any later time $t = 10 + h$ by the same algebra:

$$v_{\text{av}} = \frac{(10+h)^2 - 10^2}{h} = \frac{100 + 20h + h^2 - 100}{h} = 20 + h.$$

For the distance function $s(t) = t^2$ we have the velocity function $v(t) = 2t$. This is the key computation of calculus: we compute the distance at $t + h$, subtract the distance at t and divide by h:

$$v_{\text{av}} = \frac{s(t+h) - s(t)}{h} = \frac{(t+h)^2 - t^2}{h} = \frac{t^2 + 2th + h^2 - t^2}{h} = 2t + h.$$

As h approaches zero, the average velocity for the distance function $s(t) = t^2$ approaches $v = 2t$.

5.2 Definition of the derivative

Definition

The **derivative** $f'(x)$ or df/dx is

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

The derivative might not exist. The averages $\Delta f / \Delta t$ might not approach a limit (the same one for the time running forwards and backwards). In that case $f'(t)$ is not defined.

Example 47

Calculate the instant velocity for the distance function $f(t) = t^2$:

$$\frac{\Delta f}{\Delta t} = \frac{f(t + \Delta t) - f(t)}{\Delta t} = \frac{t^2 + 2t\Delta t + (\Delta t)^2 - t^2}{\Delta t} = \frac{2t\Delta t + (\Delta t)^2}{\Delta t} = 2t + \Delta t$$

Note that we take these steps before Δt goes to zero. If we set $\Delta t \to 0$ too early, we learn nothing, as the ratio becomes $0/0$, an expression which does not have meaning so far. The theory of limits will later allow us to understand it.
The numbers Δf and Δt must approach zero together, not separately. Then, their ratio $2t + \Delta t$ gives the correct average speed.

Theorem

The derivative of the nth power is given by

$$(x^n)' = n \cdot x^{n-1}, \qquad \text{for all } n \in \mathbb{R} \setminus \{0\}$$

The exception $n = 0$ is the constant function $y = x^0 = 1$. Its derivative, as we've already discovered, is zero.

Exercise 48

Compute the derivatives: $\left(x^5\right)', \left(\frac{1}{p^3}\right)', \left(\sqrt[3]{t^2}\right)', \left(\frac{1}{\sqrt[5]{r^9}}\right)'$.

5.3 The derivatives of operations

A huge number of functions are **linear combinations** like $f(x) = x^2 + x$ or $f(x) = x^2 - x$, or $f(x) = 5x^2$ or $f(x) = x/2$. In general also all of it at once: $f(x) = 5x^2 - \frac{1}{2}x + \sqrt{3}$. You've met such linear combinations in detail in the chapter on linear algebra.
If we need to add or subtract or multiply by 5 or divide by 2, we can *do the same with the derivatives*.

Theorem

The derivative is linear, i.e., the following holds:

1. $(c \cdot f(x))' = c \cdot f'(x)$ for any constant $c \in \mathbb{R}$.
2. $(f(x) \pm g(x))' = f'(x) \pm g'(x)$.

Example 49

We show the rule for a polynomial, in our case the quadratic function $y = 3x^2 - 4x + 5$, but the rules allow any combination of f and g.

$(3x^2 - 4x + 5)' = 3(x^2)' - 4(x)' + (5)' = 3 \cdot (2x) - 4 \cdot 1 + 0 = 6x - 4.$

Exercise 50

Compute the derivatives: $\left(4x^5 - \dfrac{3}{x^6}\right)'$, $\left(\sqrt[3]{8 \cdot x^2} + \dfrac{10}{\sqrt[5]{x^9}} - 3\sqrt{4}\right)'$.

Theorem (Leibniz product rule)

For any differetiable functions f and g the following holds:

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

Example 51

Compute the derivative of the function $y = (x^2 + 4x - 6)\sqrt{x}$

$$y' = (x^2 + 4x - 6)'\sqrt{x} + (x^2 + 4x - 6)(\sqrt{x})' = (2x + 4)\sqrt{x} + (x^2 + 4x - 6) \cdot \frac{1}{2\sqrt{x}}.$$

Theorem (Quotient rule)

For any differetiable functions f and g the following holds:

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}.$$

Exercise 52

Compute the derivative of

$$y = \frac{(\sqrt{x}-3)^2}{x}$$

Theorem (Chain rule)

For any differetiable functions f and g it holds:

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

Example 53

If the function $u(x)$ has slope du/dx, determine the slope of the composed function $f(x) = (u(x))^2$.

The first observation, e.g. with $u(x) = x^2$, gives the function $f(x) = x^4$, for which $(du/dx)^2 = (2x)^2$. On the other hand $f' = (x^4)' = 4x^3$. Hence, *the derivative of* u^2 *is not* $(du/dx)^2$.
To get the correct answer, we have to start with $\Delta f = f(x + \Delta x) - f(x)$:

$$\Delta f = (u(x+\Delta x))^2 - (u(x))^2 = [u(x+\Delta x) + u(x)] \cdot [u(x+\Delta x) - u(x)]$$

due to factorization $a^2 - b^2 = (a+b)(a-b)$. Notice we don't have $(\Delta u)^2$. Now we divide the Δf, the change in u^2, by Δx

$$\frac{\Delta f}{\Delta x} = [u(x+\Delta x) + u(x)] \cdot \frac{[u(x+\Delta x) - u(x)]}{\Delta x},$$

where the second term is just du/dx. Taking the limit we get

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = 2u(x) \cdot \frac{du}{dx}.$$

Example 54

Compute the derivative of the function $f(x) = (\sqrt{x} - 1)^2$.

In agreement with the previous calculation we get:

$$f'(x) = 2 \cdot (\sqrt{x} - 1) \cdot \frac{1}{2\sqrt{x}} = 1 - \frac{1}{\sqrt{x}}.$$

Let us check the answer by computing the derivative without the rule. Factorig the square we get $(\sqrt{x} - 1)^2 = x - 2\sqrt{x} + 1$. In this form we can compute using the rule for nth power and confirm the result.

Exercise 55

Compute the derivative of the composed functions $y = (x^3 - 1)^4$ and $y = \sqrt{8 - 4x - 2x^2}$.

5.4 The derivatives of elementary functions

We state now the rules for the derivatives of exponential functions $y = a^x$ and logarithmic functions $y = \log_b(x)$, where a and b are called **base** and both $a, b \in (0,1) \cup (1, \infty)$. For the same base, the functions $y = a^x$ and $y = \log_a(x)$ are mutually inverse, so it holds

$$a^{\log_a x} = x \qquad \text{and} \qquad \log_a(a^x) = x.$$

The rules for the derivative of the functions $y = \mathrm{e}^x$ and $y = \ln(x)$, with the Euler number $\mathrm{e} = 2.71\ldots$ as a base, are particularly simple.

Theorem

For the exponential function $y = \mathrm{e}^x$ it holds

$$(\mathrm{e}^x)' = (\mathrm{e}^x).$$

For a general exponential function $y = a^x$ a factor should be added. Let us show its value using the formula of the inverse:

$$(a^x)' = \left(\mathrm{e}^{\ln(a^x)}\right)' = \left(\mathrm{e}^{x\cdot(\ln a)}\right)' = \mathrm{e}^{(\ln a)\cdot x} \cdot ((\ln a) \cdot x)' = a^x \cdot (\ln a)$$

For the composed function $y = \mathrm{e}^{u(x)}$, with an inner function $u(x)$, the chain rule gives:

$$\left(\mathrm{e}^{u(x)}\right)' = \left(\mathrm{e}^{u(x)}\right) \cdot u'(x).$$

Example 56

Compute the derivative of the functions $y = \mathrm{e}^{\sqrt{x}}$ and $y = 2^{3t+1}$.

$$\left(\mathrm{e}^{\sqrt{x}}\right)' = \mathrm{e}^{\sqrt{x}} \cdot \frac{1}{2\sqrt{x}}$$

$$\left(2^{3t+1}\right)' = 2^{3t+1} \cdot (\ln 2) \cdot 3$$

Theorem

For the logarithmic function $y = \ln(x)$ it holds

$$(\ln(x))' = \frac{1}{x}$$

Similarly as above, for a general logarithmic function $y = \log_a(x)$ a factor should be added:

$$(\log_a x)' = \left(\frac{\ln x}{\ln a}\right)' = \frac{1}{\ln a} \cdot (\ln x)' = \frac{1}{\ln a} \cdot \frac{1}{x}$$

For the composed function $y = \ln(u(x))$, with an inner function $u(x)$, the chain rule gives:

$$(\ln(u(x)))' = \frac{1}{u(x)} \cdot u'(x).$$

Example 57

Compute the derivative of the functions $y = \ln\left(x^2 + 4x + 5\right)$ and $y = \log_2(3x)$.

$$\begin{aligned}
\left(\ln\left(x^2 + 4x + 5\right)\right)' &= \frac{1}{x^2 + 4x + 5} \cdot (2x + 4) \\
(\log_2(3x))' &= \frac{1}{\ln 2} \cdot \frac{1}{3x} \cdot 3 = \frac{1}{x \cdot \ln 2}
\end{aligned}$$

Theorem

For the derivatives of sine and cosine functions holds the following:

$$(\sin(x))' = \cos(x), \qquad (\cos(x))' = -\sin(x).$$

We derive the first formula by the standard limit technique:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{h \to 0} \frac{\Delta y}{\Delta x} = \lim_{h \to 0} \frac{\sin(x + h) - \sin(x)}{h}.$$

For this we need the **addition formula** $\sin(x + h) = \sin x \cos h + \cos x \sin h$. Since we are going to look on what happens for $h \to 0$, we factor out the $\sin(x)$ and $\cos(x)$ and get

$$\lim_{h \to 0} \frac{\Delta y}{\Delta x} = \sin(x) \left(\lim_{h \to 0} \frac{\cos h - 1}{h}\right) + \cos(x) \left(\lim_{h \to 0} \frac{\sin h}{h}\right).$$

It is no longer easy to divide by h. We proceed with showing the value of the two limits without proof, which we provide in the next chapter

$$\lim_{h \to 0} \frac{\cos h - 1}{h} = 0 \qquad \text{and} \qquad \lim_{h \to 0} \frac{\sin h}{h} = 1.$$

Therefore, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sin(x) \cdot 0 + \cos(x) \cdot 1 = \cos x.$$

Example 58

Compute the derivatives of $y = 4 + \sin(2t + 1)$ and $y = \tan^2(5\omega)$.

We first note that $(\sin(u(x)))' = \cos(u(x)) \cdot u'(x)$. Therefore, it holds

$$(4 + \sin(2t + 1))' = \cos(2t + 1) \cdot 2,$$

We deduce the formula for the derivative of the tangent function from the quotient rule:

$$(\tan(x))' = \left(\frac{\sin(x)}{\cos(x)}\right)' = \frac{\cos(x) \cdot \cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos^2(x)}.$$

Therefore,

$$\left(\tan^2(5\omega)\right)' = 2 \cdot \tan(5\omega) \cdot \frac{1}{\cos^2(5\omega)} \cdot 5.$$

Exercise 59

Compute the derivative of $y = \cot(2x)$.

Theorem

For the derivatives of the inverse trigonometric functions holds the following:

$$\begin{aligned}(\arcsin(x))' &= \frac{1}{\sqrt{1-x^2}} \\ (\arccos(x))' &= -\frac{1}{\sqrt{1-x^2}} \\ (\arctan(x))' &= \frac{1}{1+x^2}\end{aligned}$$

Example 60

Compute the derivatives of

a) $y = \arcsin\sqrt{x}$

b) $y = \arctan\frac{1}{x}$

Keeping in mind the chain rule for the derivative of composed functions, we get:

a) $(\arcsin(\sqrt{x}))' = \frac{1}{\sqrt{1-(\sqrt{x})^2}} \cdot \frac{1}{2\sqrt{x}} = \frac{1}{2\sqrt{x}\sqrt{1-x}}$,

b) $\left(\arctan\left(\frac{1}{x}\right)\right)' = \frac{1}{1+(\frac{1}{x})^2} \cdot \left(\frac{-1}{x^2}\right) = \frac{-1}{x^2+1}$.

Exercise 61

Compute the derivatives of the following functions:

a) $y = 2 \cdot x^2 \cdot \sqrt{x^3} + \sqrt[3]{8x^4} - \frac{1}{2x^3} + \frac{\sqrt{2}}{2}$, for $x_0 = 1$

b) $M_{II}(x) = F \cdot (a+x) - q_2 \cdot \frac{a}{2} \cdot \left(\frac{3}{4}a + x\right) - q_2 \cdot x \cdot \frac{x}{2}$, for $x_0 = 2$

c) $y = \frac{\tan x}{\sin(2x)}$, for $x_0 = \frac{\pi}{3}$

d) $y = \cos(1) - 3 \cdot \cos^2\left(\frac{x}{3} - \frac{\pi}{6}\right)$, for $x_0 = 0$

Exercise 62

Compute the derivatives of the following functions with respect to their independent variables:

a) $N(y) = \sqrt{\dfrac{\beta + x}{1 - y}}$, for $y_0 = 0$

b) $F(u) = \dfrac{\arcsin(1 - 4u)}{2}$, for $u_0 = \frac{1}{4}$

c) $G(z) = \ln \dfrac{4}{2z - 4} - \ln 8$, for $z_0 = 3$

d) $A(t) = A_0 + 3\mathrm{e}^{-2\alpha t + t_0}$, for $t_0 = 0$

e) $V(r) = \sqrt{\dfrac{\pi p r^4}{8\eta\ell}}$, for $r_0 = \frac{1}{2}$

5.5 Overview of the necessary formulas

$$(c \cdot f(x))' = c \cdot f'(x), \text{ for any } c \in \mathbb{R},$$

$$(f(x) \pm g(x))' = f'(x) \pm g'(x),$$

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x),$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2},$$

$$(f(g(x)))' = f'(g(x)) \cdot g'(x).$$

$$(x^n)' = n \cdot x^{n-1}, \text{ for all } n \in \mathbb{R} \setminus \{0\},$$

$$(\mathrm{e}^x)' = (\mathrm{e}^x),$$

$$(\ln(x))' = \frac{1}{x},$$

$$(\sin(x))' = \cos(x),$$

$$(\cos(x))' = -\sin(x),$$

$$(\tan(x))' = \frac{1}{\cos^2(x)},$$

$$(\cot(x))' = -\frac{1}{\sin^2(x)},$$

$$(\arcsin(x))' = \frac{1}{\sqrt{1 - x^2}}$$

$$(\arccos(x))' = -\frac{1}{\sqrt{1 - x^2}}$$

$$(\arctan(x))' = \frac{1}{1 + x^2}$$

6 Applications of the derivative

6.1 The tangent line as the best linear approximation

If we focus our attention near a single point, on a very short range, a curve looks straight. Looking through microscope, or zooming in a computer program, the graph becomes nearly linear. The curve and its tangent line have the same slope at the point of tangency. A straight line can be determined by its point and slope. That is the situation with the tangent line:

1. The equation of a line has the form $y = kx + q$.
2. The number k is the slope of the line, as $\mathrm{d}y/\mathrm{d}x = k$.
3. The number q adjusts the line to go through the required point of tangency.

Theorem

The **tangent line** t to f at the point $T = [x_0, f(x_0)]$ is given by the formula:

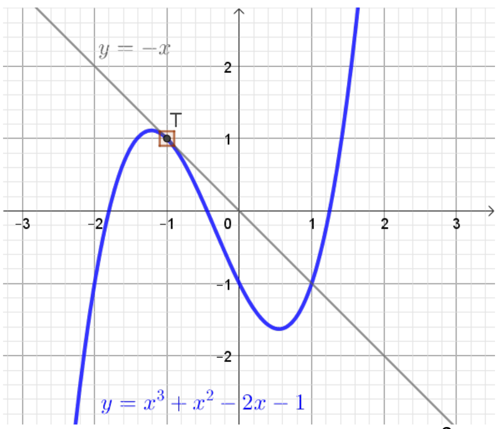
$$t : y - f(x_0) = k_t \cdot (x - x_0),$$

where $k_t = f'(x_0)$.

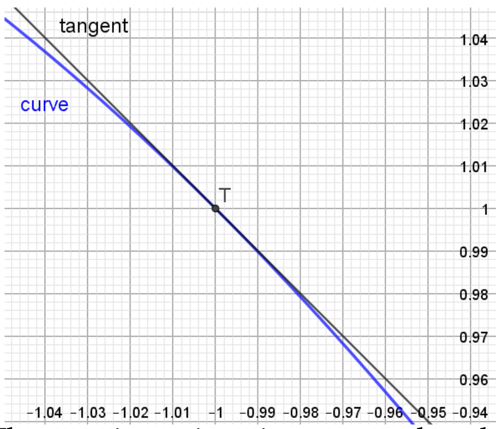
Example 63

Consider the function $f : y = x^3 + x^2 - 2x - 1$. At the point $x_0 = -1$, the value of f is $y_0 = f(-1) = 1$, which gives the point of tangency $T = [-1, 1]$. The slope of f is given by $\mathrm{d}y/\mathrm{d}x = 3x^2 + 2x - 2$. At $x = -1$ the slope is $f'(-1) = 3 \cdot (-1)^2 + 2 \cdot (-1) - 2 = -1$. The equation of the tangent line is $y - 1 = (-1)(x - (-1))$, that is

$$y = -x.$$



(a) Graph of the function $f : y = x^3 + x^2 - 2x - 1$ and its tangent at $x = -1$.



(b) The previous situation zoomed to the interval $I = \langle -1,04; -0,94 \rangle$.

Exercise 64

Write down the equation for the tangent line t to the graph of the following function:

a) $y = \dfrac{1}{(2-e^x)^2}$

b) $y = 2x \cdot \cos\left(\dfrac{x}{2}\right) + 1$

c) $y = e^{\sin\left(\frac{x}{2}\right)}$

d) $y = \sqrt{5 - e^{-4x}}$

e) $y = \cos\left(\dfrac{3x}{1-2x}\right)$

f) $y = \sqrt{1 + 2\ln(x^2 + x + 1)}$

at the point P_y, intersection point with the coordinate axis y, i.e. the line $x = 0$.
Find the slope of t.

Exercise 65

Write down the equations for the tangent lines t_i to the graph of the following function:

a) $y = \dfrac{x+1}{x^2+1}$

b) $y = \sqrt{x^3+1}$

c) $y = \dfrac{\ln(x^2-3)}{x}$

d) $y = \ln\left(x^2 - x + 1\right)$

e) $y = 1 - e^{x^2+2x-8}$

f) $y = x \cdot \arctan(x-2)$

at intersection points P_{x_i} with the coordinate axis x, i.e. the line $y = 0$.
Find the slopes of t_i.

There is another important line, closely connected to the tangent line.

Definition

The line perpendicular to the tangent and to the curve passing through the point of tangency is called **normal line**. It is usually denoted by n.

Let us discuss its slope. According to the rule that slopes of perpendicular lines multiply to give -1, the following holds:

Theorem

If the tangent has slope m, the normal line has slope $-1/m$.

Example 66

Determine the tangent and normal line to the curve $y = x^3 - 2$ at the point of tangency $[2; 6]$.

The slope of the tangent line is

$$k_t = y'(2) = \left(3x^2\right)\Big|_{x=2} = 12.$$

Hence, the point-slope equation of the tangent line is

$$t: \quad y - 6 = 12(x-2).$$

As $k_n = -1/k_t$ the point-slope equation of the normal line is

$$n: \quad y - 6 = \frac{-1}{12}(x-2).$$

6.2 Even better approximation: the Taylor polynomial

In the previous section we defined a linear approximation to estimate values of a function f at a neighborhood of the point a with known value. It can be used efficiently if

- we know the value $f(a)$
- we can easily compute the value of the first derivative of f at the point a.

However, this is not always the case. For example, a linear approximation of the Euler number $\mathrm{e} \doteq 2.71\ldots$, i.e. the value of the function $y = \mathrm{e}^x$ at the point $a = 1$, is not sufficient for precise calculations.
We can't use the tangent line at the point $a = 1$, as the value of the function and the derivative is just e.
Hence, we are forced to use the tangent approximation at $a = 0$, which gives us

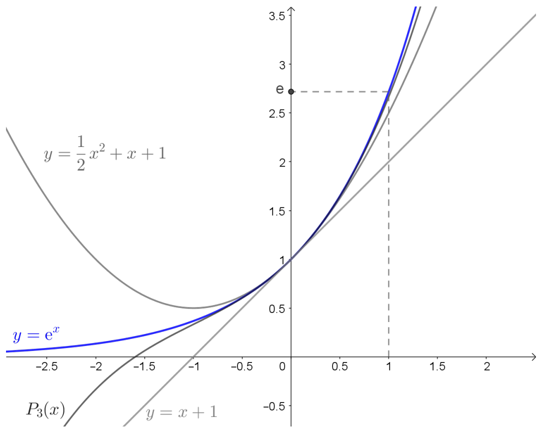
$$\mathrm{e}^1 \approx \mathrm{e}^0 + \mathrm{e}^0(x-0) = 1 + x = 2,$$

which is far from satisfactory.
It is a straightforward idea to approximate the value with a quadratic function, also with the help of the second derivative, which gives us:

$$\mathrm{e} \approx \mathrm{e}^0 + \mathrm{e}^0(x-0) + \frac{\mathrm{e}^0}{2!}(x-0)^2 = \left(1 + x + \frac{x^2}{2}\right)_{x=1} = 2.5.$$

The higher degree polynomials give us required accuracy:

$$\begin{aligned}
\mathrm{e} &\approx \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!}\right)_{x=1} &&= 2 + \tfrac{2}{3} = 2.\bar{6}, \\
&\approx \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!}\right)_{x=1} &&= 2 + \tfrac{17}{24} = 2.708\bar{3}, \\
&\approx \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!}\right)_{x=1} &&= 2 + \tfrac{43}{60} = 2.71\bar{6},
\end{aligned}$$



Definition

The polynomial of the degree $n \in \mathbb{N}$ of the form

$$T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

is called the **Taylor polynomial** of the order n for the function f centred at the point a.

We can use the Taylor polynomial to approximate the values of f at a neighborhood of a if

- we know the value $f(x_0)$,
- we can easily compute the first n derivatives of the function f at the point a.

For the centre at the origin, $a = 0$, the polynomial is also called **Maclaurin polynomial** and denoted by $M_n(x)$.

Find a polynomial that approximates the given function f in the neighbourhood of the point $x \in D(f)$ with the smallest possible error.

Theorem (Taylor)

Let f be a function with continuous derivatives up to the order $n+1$ in some neighborhood $N(a)$ of the point $a \in D(f)$. Then the following holds

$$f(x) = T_n(x) + R_{n+1}(x),$$

on $N(a)$, with the remainder term $R_{n+1}(x)$ of the form

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \qquad \text{with } \xi \in N(a).$$

From the remainder term we can estimate the error caused by taking $T_n(x)$ instead of $f(x)$.

Example 67

Estimate the error of taking $\mathrm{e} \approx T_5(1) = 2.71\bar{6}$.

Using the Taylor theorem, we compute the value of the remainder:

$$R_6(1) = \frac{f^{(6)}(\xi)}{6!}(x)^6 = \frac{\mathrm{e}^{\xi}}{6!}1^6 \leq \frac{2.72}{6!} \leq 0.0038 = \mathcal{O}(10^{-3}).$$

The first inequality follows from the fact that $\xi \in (0,1)$ and due to monotonicity of $y = \mathrm{e}^x$ we can majorize the error by taking $\xi = 1$, i.e. $f^{(6)}(\xi) = \mathrm{e}^1$. The second inequality results just from rounding up.
The efficiency of the approximation depends heavily on the magnitude of the factors in the remainder term. The error is "small" if

- $(x-a)$ is small, i.e. x is close to a,
- $n!$ is large, i.e. the order n is large,
- $|f^{(n+1)}(x)|$ is numerically small in $N(a)$.

However, the form of the remainder term enables us to compute the error, or at least its order of magnitude.

Example 68

Estimate the value of $\sin\left(\frac{1}{2}\right)$ correctly up to 5 decimal places.

We take $y = \sin(x)$ as the function for our approximation. To show the importance of $(x-a)$ being small we compute, for comparison reasons, its Maclaurin polynominal and Taylor polynomial at $a = \frac{\pi}{6}$.
Let us compute the first five derivatives of $\sin(x)$ and their values at $a = 0, \frac{\pi}{6}$. We get:

$$\begin{array}{lll}
f(x) = \sin(x) & f(0) = 0 & f(\frac{\pi}{6}) = \frac{1}{2} \\
f'(x) = \cos(x) & f'(0) = 1 & f'(\frac{\pi}{6}) = \frac{\sqrt{3}}{2} \\
f''(x) = -\sin(x) & f''(0) = 0 & f''(\frac{\pi}{6}) = -\frac{1}{2} \\
f'''(x) = -\cos(x) & f'''(0) = -1 & f'''(\frac{\pi}{6}) = -\frac{\sqrt{3}}{2} \\
f^{(4)}(x) = \sin(x) & f^{(4)}(0) = 0 & f^{(4)}(\frac{\pi}{6}) = \frac{1}{2} \\
f^{(5)}(x) = \cos(x) & f^{(5)}(0) = 1 & f^{(5)}(\frac{\pi}{6}) = \frac{\sqrt{3}}{2}
\end{array}$$

Substituting these values into the Maclaurin and Taylor formula we get:

$$\begin{aligned}
M_5(x) &= 0 + 1 \cdot x + \frac{0}{2!}x^2 - \frac{1}{3!}x^3 + \frac{0}{4!}x^4 + \frac{1}{5!}x^5 \\
&= x - \frac{1}{6}x^3 + \frac{1}{120}x^5. \\
T_5(x) &= \frac{1}{2} + \frac{\sqrt{3}}{2} \cdot \left(x - \frac{\pi}{6}\right) - \frac{1}{4}\left(x - \frac{\pi}{6}\right)^2 - \frac{\sqrt{3}}{12}\left(x - \frac{\pi}{6}\right)^3 + \\
&\quad + \frac{1}{48}\left(x - \frac{\pi}{6}\right)^4 + \frac{\sqrt{3}}{240}\left(x - \frac{\pi}{6}\right)^5.
\end{aligned}$$

The value of Maclaurin expansion at $x = \frac{1}{2}$ is

$$T_5\left(\frac{1}{2}\right) = \frac{1}{2} - \frac{1}{48} + \frac{1}{3840} = \frac{1841}{3840} = 0.4794\ldots$$

and the error, estimated using the formula for the remainder,

$$R_6\left(\frac{1}{2}\right) = \frac{|-\sin(\xi)|}{(6)!}\left(\frac{1}{2}\right)^6 \leq \frac{|\sin\frac{\pi}{6}|}{720}\cdot\frac{1}{64} \doteq 0.00001085\ldots = \mathcal{O}(10^{-5}).$$

Hence, we have achieved the desired precision. Note that for $\frac{\pi}{6} > \frac{1}{2}$ the value $\sin\frac{\pi}{6} > \sin(\xi)$ for any $\xi \in (0, \frac{1}{2})$ due to the monotonicity of $\sin(x)$ on the interval $(0, \frac{\pi}{6})$.

We proceed with the Taylor formula at $x = \frac{1}{2}$ and estimate the error from the remainder:

$$R_6\left(\frac{1}{2}\right) = \frac{|-\sin(\xi)|}{(6)!}\left(\frac{1}{2} - \frac{\pi}{6}\right)^6 \leq \frac{\frac{1}{2}}{720}\cdot\left(\frac{3-\pi}{6}\right)^6 = \mathcal{O}(10^{-9}).$$

It is clear that due to the fact that $(x - a)$ is smaller we obtained much better result (or the same precision much more quickly).

Exercise 69

Write down the Maclaurin polynomial of the order 4 for the function

$$\text{a) } f : y = x^2 e^x, \qquad \text{b) } g : y = e^x \cdot \cos(x).$$

Exercise 70

Estimate the values of $\sin(1^\circ)$, $\sin(1)$, $\tan(1)$ and $\arctan(1)$ correctly up to 3 decimal places.

6.3 Minimum and maximum in applications

Definition

We say that the function f has a **local maximum** at the point $x_0 \in D(f)$ if there exists a punctured neighborhood $N(x_0)$ of the point x_0 such that

$$f(x) < f(x_0) \qquad \text{for all } x \in N(x_0).$$

Similarly, the function f has a **local minimum** at the point $x_0 \in D(f)$ if

$$f(x) > f(x_0) \qquad \text{for all } x \in N(x_0)$$

How do you identify maximum or minimum?
Typically, the slope is zero. *If* $\mathrm{d}f/\mathrm{d}x$ *exists, it must be zero.* The tangent line is horizontal. The graph changes from increasing to decreasing. The slope changes from positive to negative. This turning point of f' is called **a stationary point**.

It is also possible that the graph has a corner, and thus no derivative. These points are called **rough points**.

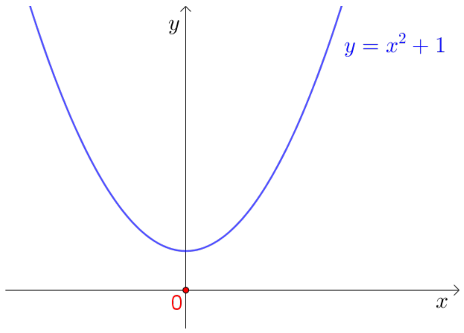
Last but not least we should check the **endpoints** of the domain.

Theorem (Fermat's theorem)

If f is differentiable at $x_0 \in (a, b)$, and $f'(x_0) \neq 0$, then x_0 is not a local extremum of f.

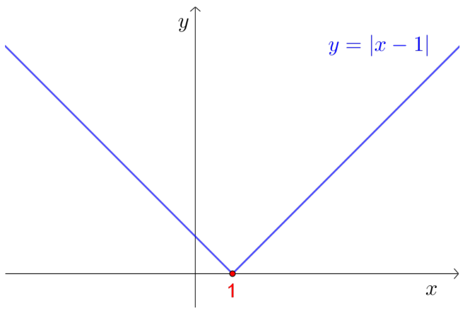
Fermat's theorem gives us a necessary condition for the existence of a local extremum. As a contrapositive statement, it allows us to rule out the points, where there is no extremum. The remaining points are co called **critical points**, suspected of the existence of an extremum. They are of the following three types:

a) **stationary points**, where $\mathrm{d}f/\mathrm{d}x = 0$,



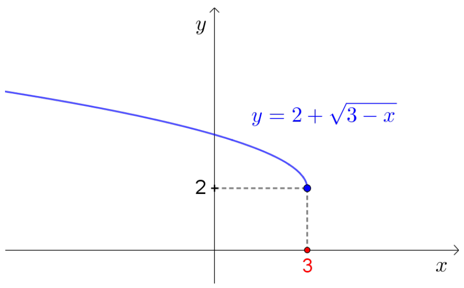
Graph of the function $y = x^2 + 1$ with the stationary point $x = 0$,

b) **rough points**, where $\mathrm{d}f/\mathrm{d}x$ does not exist,



Graph of the function $y = |x - 1|$ with the rough point $x = 1$,

c) **endpoints** of the domain,



Graph of the function $y = 2 + \sqrt{3 - x}$ with the endpoint $x = 3$.

For the most common case of the stationary points there is an easy-to-check sufficient condition for the existence of a local extremum.

Theorem (Second derivative test)

If $f'(x_0) = 0$ and $f''(x_0) > 0$, then the function f has a local minimum at the point x_0.

If $f'(x_0) = 0$ and $f''(x_0) < 0$, then the function f has a local maximum at the point x_0.

To find a maximum or minimum, we just find critical points of f. We solve the equation $f'(x) = 0$ and then check the rough points, where $f'(x)$ doesn't exist, and endpoints. The idea is clear, but to be honest, that is not where the problem starts.

In reality, the first (and often the hardest) step is to choose the unknown variable, which should be minimized, and find the function describing its behaviour. We and only we ourselves decide what will be x and how would $f(x)$ look like. The equation $\mathrm{d}f/\mathrm{d}x = 0$ comes out by a standard procedure, often easily with the help of computer. On the other hand no computer so far is able to analyse the situation and propose the correct form of f.

The heart of this subject is in word problems. The procedure of solving the problem can be divided into steps:

1. Find (propose) the quantity x to be minimized or maximized.
2. Express the quantity x as a function $f(x)$.
3. Compute $f'(x)$ and solve $f'(x) = 0$.
4. Check all critical points for $f_{\min}$ and $f_{\max}$.

Example 71

Barrel problem.
The army is looking for a big amount of 500 litre barrels for fuel. Due to the shortage of the metal plates, you should propose a shape of the barrels (the radius r and the height h) spending the least of the worthy material.

The volume of the barrel is obtained from the formula for the cylinder

$$V = \pi r^2 h = 500\,\mathrm{dm}^3.$$

The surface of the cylindrical barrel consists of two circles (the bottom and the top) and the rectangular body. Its area A is their sum:

$$A = 2(\pi r^2) + 2\pi r h.$$

The function A should be made minimal. However, it depends on the two unknowns r and h and we need to minimize function of one variable only. The two variables are connected through the formula for V, which gives us the possibility of expressing the height h in terms of r. Indeed,

$$h = \frac{500}{\pi r^2}.$$

In this way we obtain the formula for A as a function of one variable r:

$$A(r) = 2(\pi r^2) + 2\pi r \frac{500}{\pi r^2} = 2(\pi r^2) + \frac{1000}{r},$$

with the domain $D(A) = (0, \infty)$.
The rest is pretty standard. To find the minimum of the function $A(r)$ we first compute the derivative

$$\frac{\mathrm{d}A}{\mathrm{d}r} = 4\pi r + 1000 \cdot \frac{-1}{r^2}$$

and use the Fermat theorem $dA/dr = 0$ to compute the stationary points. This equation has unique solution

$$r_0 = \sqrt[3]{\frac{1000}{4\pi}} \doteq 4.301\,\mathrm{dm}.$$

We can check that this is the minimum using the second derivative test. Indeed,

$$\frac{\mathrm{d}^2 A}{\mathrm{d}r^2} = 4\pi + \frac{2000}{r^3} > 0.$$

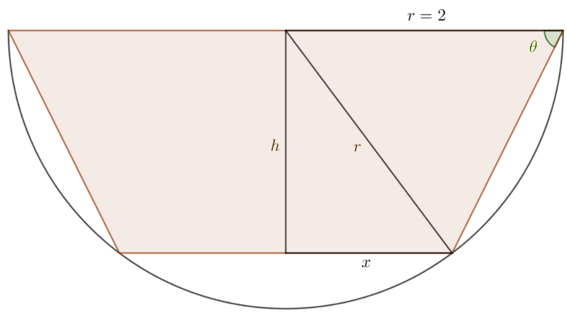
Remark

Notice that $h = 2r$, which means that height is equal to the diameter of the barrel. This is another manifestation of a symmetry so often seen in the minimization problems.

Exercise 72

Drainage canal.
The company building a drainage canals should dredge a semicircle thalwegs of radius 2 m. The canals should be concreted into the form of a trapezoid with the bottom parallel to the surface, see figure below. Find the shape of the trapezoid so that is allows maximal possible streaming (in that case the trapezoid has maximal area).

Sectional view of the drainage canal is as follows:



Exercise 73

Baywatch.
You are standing near the side of a large wading pool of uniform depth when you see a child in trouble. You can run at a speed $v_1 = 7.1\,\mathrm{m/s}$ on land and swim at the speed $v_2 = 1.6\,\mathrm{m/s}$ in the water. Your perpendicular distance from the side of the pool is a, the child's perpendicular distance is b, and the distance along the side of the pool between the closest point to you and the closest point to the child is c (see the figure below). Without stopping to do any calculus, you instinctively choose the quickest route (shown in the figure) and save the child. Our purpose is to derive a relation between the angle θ_1 your path makes with the perpendicular to the side of the pool when you're on land, and the angle θ_2 your path makes with the perpendicular when you're in the water. To

do this, let x be the distance between the closest point to you at the side of the pool and the point where you enter the water. Write the total time you run (on land and in the water) in terms of x and find its minimum.

Exercise 74

Fencing a pasturage.
A rancher needs to fence a rectangular pasturage area next to a straight river, using 1200 m of fencing. The side next to the river will not be fenced, to allow the cattle drinking and freshening up in the river. Advice the rancher the dimensions of the rectagnle to maximize the area of the pasturage. What is the maximum area?

Exercise 75

Running a hotel.
A new 120-room hotel to be opened in Prague is setting up its prices. The manager knows that they will rent all of its rooms if they charge €50 per room and for each €2 increase per room, three fewer rooms will be rented per night. What rent per room would maximize the profit per night?

Exercise 76

Cutting a beam.
The strength of a rectangular beam is given by $S = \nu \cdot w \cdot d^2$, with width w and depth d. Find the dimensions of the strongest beam that can be cut from a cylindrical log of larch wood ($\nu = 0.35$) of radius $r = 30\,\text{cm}$.

Exercise 77

Shipping a parcel to the USA.
The U.S. post office will accept a box for shipment only if the sum of the length and girth (distance around) is at most 274 cm. Find the dimensions of the largest acceptable box with square front and back.

6.4 Properties of the grahp of an elementary function

Monotonicity

Suppose that df/dx is positive at a point x_0. Then the tangent line slopes upward. Therefore it is increasing (as a linear function) and the function $f(x)$ itself is also increasing at the point x_0.

Theorem

If $f'(x_0) > 0$, then the function f is strictly increasing at the point x_0.

If $f'(x_0) < 0$, then the function f is strictly decreasing at the point x_0.

This "local" theorem describing behaviour can be easily generalized to "global" open intervals:

Theorem

If $f'(x_0) > 0$ for all $x \in I = (a, b)$, then the function f is strictly increasing on the interval I.

If $f'(x_0) < 0$ for all $x \in I = (a, b)$, then the function f is strictly decreasing on the interval I.

Remark

Note that the preceding theorem cannot be generalized on the union of intervals.

Example 78

Find the intervals of the strict monotonicity of the function

$$y = x^2 - 12\ln(x - 1).$$

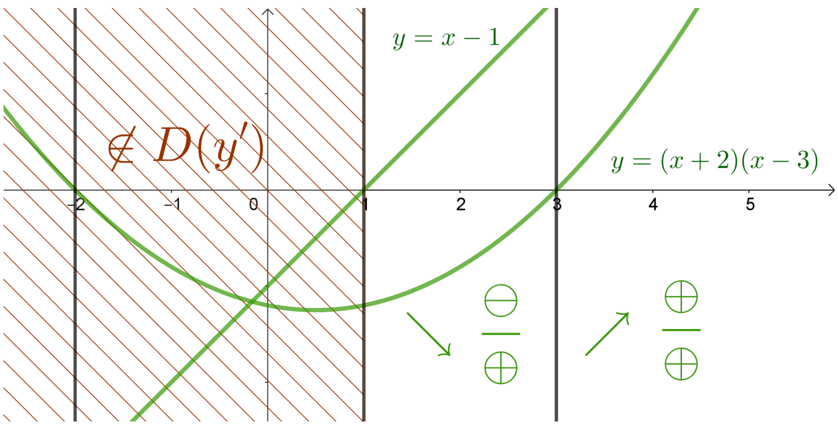
We start with computing the domain. Here, we have only one condition, required by the definition of the logarithmic function, namely $x - 1 > 0$. Therefore, the domain is $D = (1, \infty)$.
We proceed with the first derivative,

$$y' = 2x - 12 \cdot \frac{1}{x-1}.$$

Note that $D_{y'} = D = (1, \infty)$, even if the function $2x - 12/(x - 1)$ has larger domain.
As we need to solve two inequalities, $y' > 0$ and $y' < 0$, we find the zero points of y'.

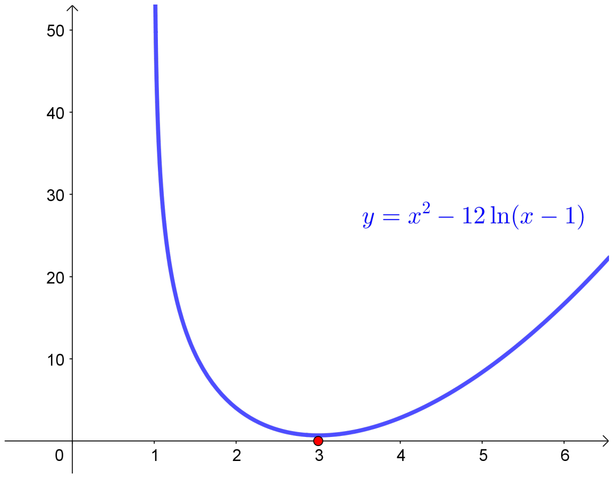
$$y' = 2x - 12 \cdot \frac{1}{x-1} = \frac{2x(x-1) - 12}{x-1} = \frac{2x^2 - 2x - 12}{x-1} = \frac{2(x+2)(x-3)}{x-1} = 0.$$

This rational expression has three roots, $x_i = -2, 1, 3$, which divide the domain to the subintervals, where y' does not change its sign and therefore it will be possible to decide which inequality is fulfilled on a particular interval.
As the roots $x_1 = -2$, $x_2 = 1$ do not belong to the domain, we have just single root $x = 3$ and we solve the inequalities by the graphical method on the intervals $(1, 3)$ and $(3, \infty)$.
The denominator is positive in both cases, so the sign of the numerator decides on the result:



Intervals of monotonicity.

The function is increasing on the interval $(3, \infty)$ and decreasing on $(1, 3)$.



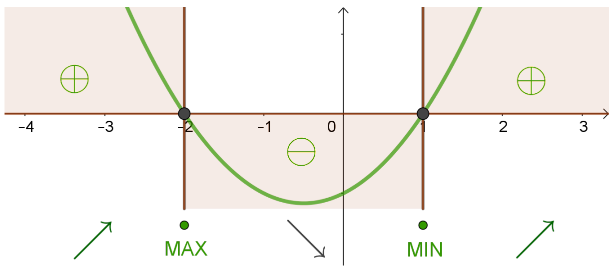
Intervals of monotonicity.

Example 79

Find intervals of monotonicity and all local extremes of the function

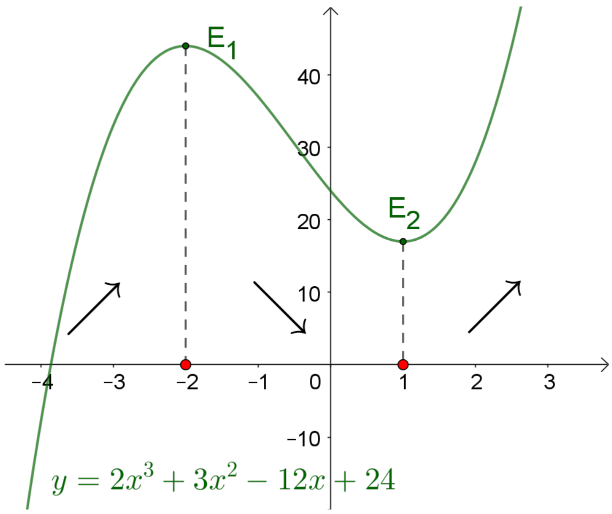
$$y = 2x^3 + 3x^2 - 12x + 24.$$

We start with checking the domain, here it is easy, $D = \mathbb{R}$. We proceed with the first derivative, $y' = 6x^2 + 6x - 12$.
We solve both tasks at once. As we need to solve the the equation $y' = 0$ and the two inequalities, $y' > 0$ and $y' < 0$, we first find the zero points of y', which divide the domain to the subintervals, where y' does not change its sign and therefore it is easy to decide which inequality is fulfilled on a particular interval.
We rewrite the equation $y' = 0$ to the form $6(x+2)(x-1) = 0$. Hence, the stationary points are $x_1 = -2$, $x_2 = 1$, we have no rough points and no endpoints. We easily solve the inequalities by the graphical method:



Intervals of monotonicity.

The function is increasing on the intervals $(-\infty, -2)$ and $(1, \infty)$ (however not on their union!) and decreasing on $(-2, 1)$.
Therefore, there is a local maximum at $x = -2$ and a local minimum at $x = 1$. Note that we don't even need the second derivative test in this case.



Graph of the function with marked intervals of monotonicity and local extremes.

Exercise 80

For the given functions:

a) $y = x + \ln\left(2x^2 - x + 1\right)$

b) $y = \ln\sqrt[3]{x^2} + x$

c) $y = \dfrac{\ln x}{1 - \ln x}$

d) $y = \ln\left(\dfrac{1-x}{x+2}\right)$

e) $y = \dfrac{1}{x \cdot \ln x}$

f) $y = \ln\left(x^2\right) - x^2$

g) $y = \sqrt{\dfrac{x-2}{3-x}}$

h) $y = \arctan\left(x + \dfrac{1}{x}\right)$

i) $y = \mathrm{e}^{\cos(2x)}$

j) $y = 1 + 2\sin^3 x$

k) $y = (1 + 2\sin x)^3$

l) $y = \sqrt{(x-1)\cdot(x+1)\cdot(x+3)}$

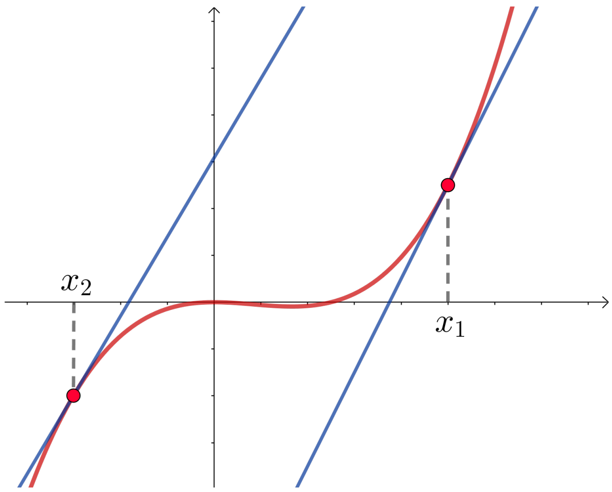
m) $y = \dfrac{x^3}{x^2 + 4x + 3}$

n) $y = \dfrac{x^2+4}{x^2 - 3x + 4}$

find the intervals, where the function is increasing (respectively decreasing). Compute the coordinates of the local maxima and minima.

Convexity as an expression of a curvature of the graph

The curvature of the graph can be also described by the tangent line. If the *tangent line* to the graph of the function $y = f(x)$ at the point $[x_0, f(x_0)]$ *is lying below* the graph of the function at some neigborhood of x_0, we call the function **convex** at the point (x_0). Similarly, if the tangent line is lying above the graph, it is **concave** at the point (x_0).



A function convex at x_1 and concave in x_2.

These considerations are not convenient for practical calculations, therefore we have the following easy criterion.

Theorem

If $f''(x_0) > 0$, then the function f is convex at the point x_0.

If $f''(x_0) < 0$, then the function f is concave (convex negative) at the point x_0.

The theorem extends very easily to open intervals:

Theorem

If $f''(x_0) > 0$ for all $x \in (a, b)$, then the function f is convex on the interval (a, b).

If $f''(x_0) < 0$ for all $x \in (a, b)$, then the function f is concave on the interval (a, b).

The latter theorem

1. geometrically justifies our second derivative test,
2. cannot be generalized on the union of intervals.

Example 81

Find the intervals of the convexity for the function

$$y = x^4 - 4x^3.$$

We start with checking the domain, $D = \mathbb{R}$. Next, the first and second derivative are

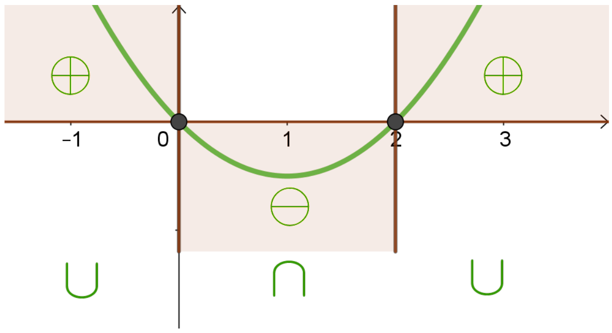
$$\begin{aligned} y' &= 4x^3 - 12x^2, \\ y'' &= 12x^2 - 24x. \end{aligned}$$

Similarly as by the monotonicity, we need to solve two inequalities, only here with the second derivatives, $y'' > 0$ and $y'' < 0$.

Again, we first find the zero points of y''. Factoring out the common multiple 12 we rewrite the equation $y'' = 0$ to the form

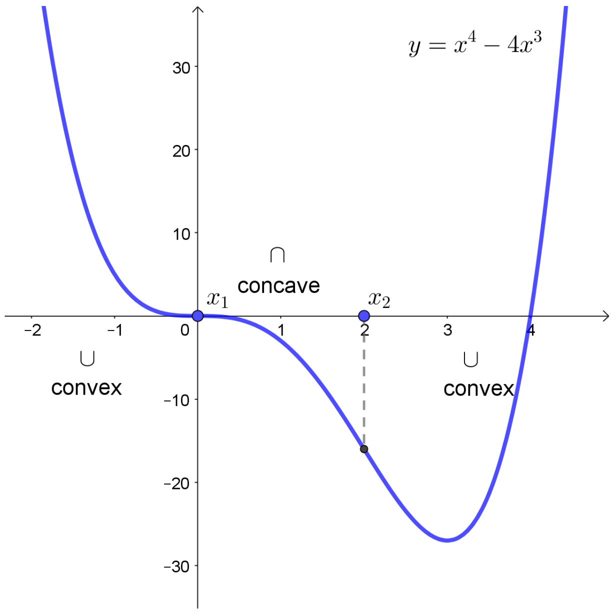
$$12x(x-2) = 0.$$

So the stationary points are $x_1 = 0,\ x_2 = 2$, with no rough points and no endpoints. Again, We easily solve the inequalities by the graphical method:



Intervals of convexity and concavity. The symbol $\oplus$ denotes the interval, where $y'' > 0$, the symbol $\ominus$ denotes $y'' < 0$.

Notice that the point $x = 0$ is not an extremum, even though it holds $f'(0) = 0$.



The function $y = x^4 - 4x^3$, its intervals of convexity and concavity.

Inflection points

We describe the points on a graph, where the curvature changes of sign. In particular, it is a point where the function changes from being concave to convex, or vice versa.

Definition

We say that the point $x_0 \in D$ is the **inflection point** of the function f, if there exists a neighborhood $N_\delta(x_0)$ of the point x_0 such that

$$f''(x-\epsilon) \cdot f''(x+\epsilon) < 0 \qquad \text{for all } \epsilon \in (0,\delta).$$

This is not practical criterion that can be efficiently used by the computation. We state such a criteria in the following theorems.

Theorem (necessary condition for the existence of an inflection point)

If the function f has an inflection point x_0, then

$$f''(x_0) = 0.$$

Theorem (sufficient condition for the existence of an inflection point)

If $f''(x_0) = 0$ and $f'''(x_0) \neq 0$, then the function f has an inflection point x_0.

The thoerem can be further precised. At the inflection point x_0 the lowest non-zero derivative is of an odd order.

Example 82

Find all inflection points of the function $y = \sin(2x)$.

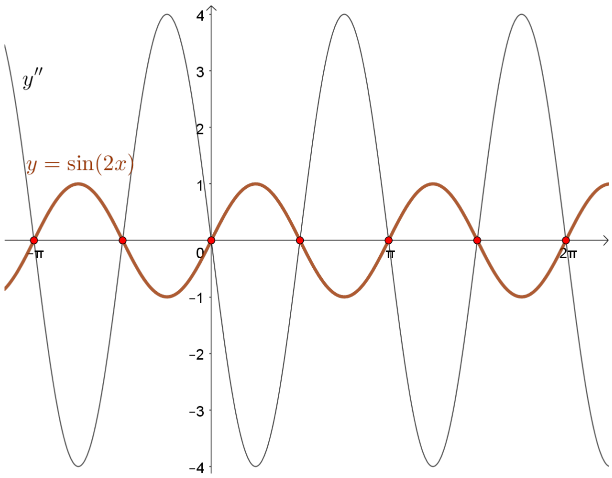
We compute the second derivative y''

$$\begin{aligned} y' &= \cos(2x) \cdot 2, \\ y'' &= 2(-1)\sin(2x) \cdot 2 = (-4)\sin(2x) \end{aligned}$$

and set $y'' = 0$. We see that the inflection points coincide with the intersection point of y with the x axis,

$$x_i = \left\{ 0 + k\frac{\pi}{2} \,\middle|\, k \in \mathbb{Z} \right\}.$$

Moreover, if $y = \sin(2x)$ is positive, then it is concave, and if y is negative, it is convex.



The intervals of convexity and concavity of the function $y = \sin(2x)$.

Exercise 83

For the given functions:

a) $y = \ln(x) - \sqrt{x}$

b) $y = \ln\left(x^2 - 1\right)$

c) $y = \ln\left(\frac{1}{\sqrt{x}}\right) - x^2$

d) $y = \ln(x-1) + \frac{x^2}{2}$

e) $y = \left(1 - \frac{1}{x}\right)^3$

f) $y = \sqrt{x^2 - 1}$

g) $y = \frac{e^x}{x}$

h) $y = x^4 \cdot e^x$

i) $y = \sqrt{e^{1-x}}$

j) $y = e^{x^2-1}$

k) $y = \frac{2 - x^2}{e^x}$

l) $y = (x^2 - 2) \cdot e^{x-1}$

m) $y = \sin(x) \cdot e^x$

n) $y = e^{\frac{2}{1-x}}$

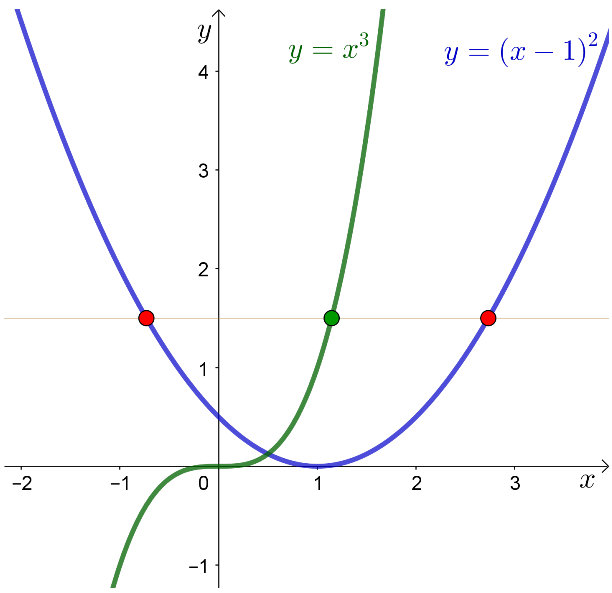
find the point of inflection. Next, determine the intervals, where the function is convex (respectively concave).

Transposing formulae

The formula $y = f(x)$ describes evolution of physical quantity f depending on another physical quantity x. In physics, some processes are idealised as *reversible*. In this case, the function describing the value quantity f is uniquely determined by the value of x. We say that f is **one-to-one**.
This property can be formulated geometrically as follows:

Definition

The function $y = f(x)$ is called **one-to-one** if and only if the graph of f and a horizontal line $y = q$ have at most one intersection point for any $q \in \mathbb{R}$.



The prototypes of one-to one functions are odd powers, e.g. $y = x^3$. On the other hand, the even powers are not one-to-one.

There is an easy-to-check criterion for any f to be one-to-one.

Theorem

- If the function f is increasing, it is one-to-one.
- If the function f is decreasing, it is one-to-one.

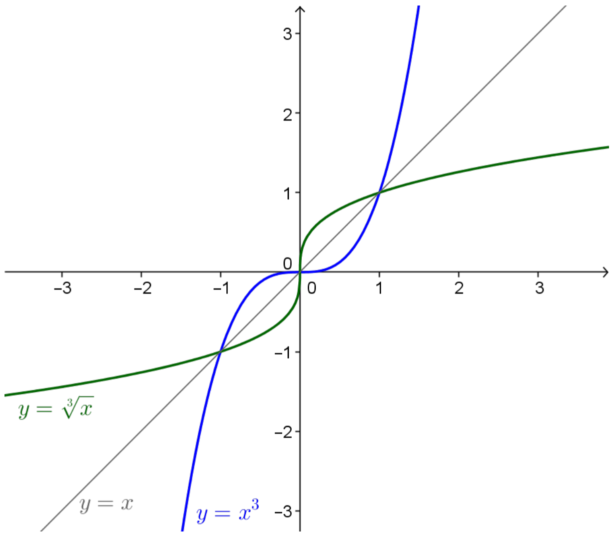
The function describing the process returning the system into the original state can be computed by transposing the formula for f. The resulting function is called **the inverse function to** f and denoted by f^{-1}.

Theorem

For the one-to-one function $y = f(x)$ with the domain D_f and range I_f, there exists unique one-to-one inverse function to f defined on $I_f = D_{f^{-1}}$ by the formula

$$f^{-1}(y) = x \qquad \text{if and only if} \qquad f(x) = y.$$

Moreover, if f^{-1} is the inverse function to f, then f is the inverse function to f^{-1}; i.e. the inverse relation is mutual. Therefore, the graphs of the mutually inverse functions are axially symmetric with respect to the line (axis) $y = x$.



The graphs of the mutually symmetric functions are axially symmetric.

Theorem

- If f is increasing then f^{-1} is increasing.
- If f is decreasing then f^{-1} is decreasing.

Note that the notation of the inverse function is ambiguous. In this context, the superscript *does not* mean "f to the power of -1", as the inverse is made with respect to the composition of functions and not multiplication. Therefore, it holds

$$f^{-1}(f(x)) = x$$

and *not* $f^{-1}(x) \cdot f(x) = 1$. In another words, the inverse function f^{-1} is *not* equal to the reciprocal function $y = 1/f$, i.e. $f^{-1}(x) \neq 1/f(x)$.

What is the relation between the domain and range of the mutually inverse functions?
According to the definition and properties of f^{-1} it holds:

- $D_f = I_{f^{-1}}$ and $I_f = D_{f^{-1}}$,
- for every $x \in D_f$ and $y \in D_{f^{-1}}$ it holds

$$f^{-1}(f(x)) = x \qquad \text{and} \qquad f(f^{-1}(y)) = y.$$

We often need to transpose a formula, which is not one-to-one, e.g. for the function $y = x^2$. We now describe how this could be done.

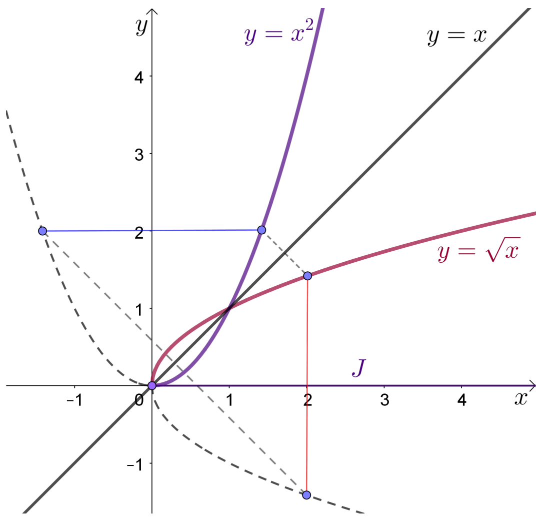
The procedure of finding the inverse function $y = f^{-1}(x)$ **to the function** $y = f(x)$ can be described as follows:

1. find the domain D_f.
2. Decide if the function f is one-to-one. If it is not, find the biggest interval J, where f is one-to-one, and take J as the domain of f, make a **restriction** of f on J. It is denoted by $f\,|_J$.
3. Compute the transposed formula for its inverse function.
4. Compute the domain and the range of the inverse function.

Example 84

Construct an inverse function to the function $f : y = x^2$.

The function f is even, hence not one-to one. We choose $J = \langle 0, \infty)$ and construct f^{-1} for $f\,|_J$. As on a larger interval f is not one-to-one, the resulting f^{-1} would not be a function.



Constructing an inverse function f^{-1} to the even function, $y = x^2$.

Notice that there might be a freedom of choice in the decision about the interval J. In the preceding example, we could choose $J = (-\infty, 0\rangle$. In practise, we make the decision based on which points $x \in D_f$ we want to map.

Example 85

Let be given the function

$$f(x) : y = 1 - \ln\left(-1 + \sqrt{x}\right).$$

Decide if f is one-to-one. Compute the transposed formula for its inverse function f^{-1}. Write down the domains D_f, $D_{f^{-1}}$ and ranges I_f, $I_{f^{-1}}$.

We start with the domain of f. From the two conditions involved by the square root, $x \geq 0$, and the logarithm, $-1 + \sqrt{x} > 0$, we get $D_f = (1, \infty)$.
Next we check if the function f is one-to-one, based on its monotonicity. We compute f',

$$y' = -\frac{1}{\sqrt{x} - 1} \cdot \frac{1}{2\sqrt{x}}.$$

This expression is negative on D_f, as the first fraction is positive for $x > 1$, i.e. exactly on D_f, and the second fraction is positive even on the bigger interval $\langle 0, \infty)$. Therefore, f is decreasing, thus one-to-one.
We compute the formula for the inverse by switching $x \leftrightarrow y$ and expressing y:

$$\begin{aligned} x &= 1 - \ln\left(-1 + \sqrt{y}\right) \\ \ln\left(-1 + \sqrt{y}\right) &= 1 - x \\ -1 + \sqrt{y} &= \mathrm{e}^{1-x} \\ \sqrt{y} &= 1 + \mathrm{e}^{1-x} \\ y &= \left(1 + \mathrm{e}^{1-x}\right)^2 \end{aligned}$$

The domain $D_{f^{-1}} = \mathbb{R} = I_f$ and the range $I_{f^{-1}} = (1, \infty) = D_f$, as expected.

Example 86

Let be given the function

$$f(x) : y = 3 - \frac{2}{1 + 2x + x^2}.$$

Decide if f is one-to-one. If it is not, find the biggest interval J, where f *is* one-to-one and take $f\,|_J$.
Next, compute the transposed formula for its inverse function f^{-1}. Write down the domains D_f, $D_{f^{-1}}$ and ranges I_f, $I_{f^{-1}}$.

We first rewrite the formula as follows:

$$f : y = 3 - \frac{2}{(1 + x)^2}.$$

Then, we see the domain $D = \mathbb{R} \setminus \{-1\} = (-\infty, -1) \cup (-1, \infty)$ more easily.
In order to show that f is one to one, we show the intervals of monotonicity. The derivative

$$y' = (-2)(-2)(1 + x)^{-3} = \frac{4}{(1 + x)^3}$$

is positive on $J_1 = (-1, \infty)$ and negative on $J_2 = (-\infty, -1)$. Therefore, f is not monotone. However, if we choose just one of the two intervals, f would be monotone, therefore one-to-one. It can be also checked by the graph.
We choose J_1 and restrict f on this interval, i.e. take J_1 as the domain of f. Therefore, f is increasing on J_1, hence one-to-one.
We can also compute the inverse. We do it by switching $x \leftrightarrow y$ in the formula for f and expressing y:

$$\begin{aligned} x &= 3 - \frac{2}{(1+y)^2} \\ \frac{2}{(1+y)^2} &= 3 - x \\ (1+y)^2 &= \frac{2}{3-x} \\ 1+y &= \sqrt{\frac{2}{3-x}} \\ y &= \sqrt{\frac{2}{3-x}} - 1 \end{aligned}$$

Just note that we have taken the *positive* square root in the next-to-last line, in order that $x \overset{f}{\to} y \overset{f^{-1}}{\to} x$. Therefore, we get the range $I_{f^{-1}} = (-1, \infty) = J_1$, i.e., the interval we chose as D_f. The domain, $D_{f^{-1}} = (-\infty, 3)$, coincides with the range I_f.

Exercise 87

Decide if the given functions:

a) $f(x) : y = 1 - \ln\left(-1 + \sqrt{x}\right)$

b) $M(R) : M = \pi(R^4 - r^4)$

c) $f(x) : y = \ln(1 - e^x)$

d) $p(b) : \frac{p}{q} = \sqrt{\frac{a+2b}{a-2b}}$

e) $f(x) : y = 3 + 2 \cdot \arccos\frac{x}{2}$

f) $S(L) : S = \sqrt{\frac{3d(L-d)}{8}}$

g) $f(x) : y = \frac{4x-1}{x+3}$

h) $y(r) : y + x = \frac{r}{4+r}$

i) $f(x) : y = \ln(x-1) - \ln(x+1)$

j) $f(x) : y = 3 - \frac{2}{1+2x+x^2}$

k) $f(x) : y = \frac{1}{\sin(x)}$

are one-to-one. Compute the transposed formula for its inverse function f^{-1}. Write down their domains and ranges.

6.5 Efficient computation of the limits with the l'Hôpital rule

When two functions approach zero, their ratio might do anything. That is why we call $0/0$ an **indeterminate expression**. The results depends on the particular form of the expression hidden behind the zeros in the numerator and the denominator. We might have

$$\lim_{h\to 0} \frac{h^2}{h} = 0 \quad \text{or} \quad \lim_{h\to 0} \frac{h}{h} = 1 \quad \text{or} \quad \lim_{h\to 0} \frac{7h}{h} = 7 \quad \text{or} \quad \lim_{h\to 0} \frac{\sqrt{h}}{h} = \infty.$$

What only matters is *whether the numerator goes to zero more quickly than denominator*.
There are eight typical indeterminate expressions:

$$\frac{0}{0}, \quad \frac{\pm\infty}{\pm\infty}, \quad 0\cdot\infty, \quad \infty-\infty, \quad 0^0, \quad 0^\infty, \quad \infty^0, \quad 1^\infty.$$

The efficient and powerful method to compute the limits of the indeterminate expressions is named afted Guillaume François Antoine, Marquis de l'Hôpital (1661–1704), who published it first in 1696. However the idea belongs probably to Jacob Bernoulli (1667–1748).

Theorem (L'Hôpital)

Suppose that

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$$

or

$$\lim_{x\to x_0} f(x) = \pm\infty \text{ and } \lim_{x\to x_0} g(x) = \pm\infty.$$

Then it holds

$$\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = a \qquad \Rightarrow \qquad \lim_{x\to x_0} \frac{f(x)}{g(x)} = a$$

Example 88

Compute the limits

a) $\lim_{h\to 0} \frac{\cos h - 1}{h}$ b) $\lim_{h\to 0} \frac{\sin h}{h}$

used for the deduction of the rules for the derivative of sine and cosine functions.

Both expressions are of the form $0/0$, so we can use the l'Hôpital rule.
a) Formally, we should proceed carefully and start by computing the limit of the ratio of the derivatives:

$$\lim_{h\to 0} \frac{(\cos h - 1)'}{(h)'} = \lim_{h\to 0} \frac{-\sin h}{1} = \sin(0) = 0.$$

If this limit exists, then the original limit also exists and they are the same:

$$\lim_{h\to 0} \frac{(\cos h - 1)'}{(h)'} = 0 \qquad \Rightarrow \qquad \lim_{h\to 0} \frac{\cos h - 1}{h} = 0.$$

b) In the most cases the limit exists, even if perhaps after multiple derivatives. Therefore, we simplify the formal procedure:

$$\lim_{h\to 0} \frac{\sin h}{h} \overset{\text{l'H}}{=} \lim_{h\to 0} \frac{\cos h}{1} = \cos(0) = 1.$$

Example 89

Compute the limit

$$\lim_{x\to\infty} \frac{x+\sin x}{x}.$$

The expression is of the form ∞/∞, so we can apply l'Hôpital rule:

$$\lim_{x\to\infty} \frac{x+\sin x}{x} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{1+\cos x}{1} = 1+\cos(\infty).$$

This limit does not exist, as the cosine function oscillates for $x \to \infty$. However, from this fact we cannot deduce that the original limit does not exist! Just, in this rare case, l'Hôpital rule did not give us any answer and we should proceed with a different method:

$$\lim_{x\to\infty} \frac{x+\sin x}{x} = \lim_{x\to\infty} \frac{x}{x} + \frac{\sin x}{x} = 1 + \lim_{x\to\infty} \frac{\sin x}{x} = 1,$$

because the last limit is zero, as the limit of the expression of the form n/∞, with $n \in [-1,1]$.

Exercise 90

Solve:

a) $\lim_{x\to 0} \frac{\sin(4x)}{x}$

b) $\lim_{x\to 0} \frac{\tan^2(2x)}{2x^2}$

c) $\lim_{x\to\infty} \frac{1-e^{2x}}{x^4}$

d) $\lim_{x\to\infty} \frac{\ln^2 x}{x^2+1}$

For the limits of quotient of the power functions it holds:

$$\lim_{x\to\infty} \frac{ax^m}{bx^n} = \begin{cases} \infty & \text{for } m > n \\ \frac{a}{b} & \text{for } m = n \\ 0 & \text{for } m < n \end{cases}$$

Exercise 91

Solve:

a) $\lim_{x\to\infty} \frac{4x^2+6x+9}{x^4+2x^2+1}$

b) $\lim_{x\to\infty} \frac{\sqrt[3]{8x^3-1}}{(x-2)^2}$

c) $\lim_{x\to\infty} \frac{x^3-7x^2-1}{6x^2+9x+15}$

d) $\lim_{x\to\infty} \frac{\sqrt{9x^4+8x^2+1}}{3x-1}$

e) $\lim_{x\to\infty} \frac{4x^3-1}{7x^3+6x^2+5x+4}$

f) $\lim_{x\to\infty} \frac{\sqrt{2x^4-4x^2}}{3x^2+2x+1}$

One-sided limits

Sometimes we can approach the point, where we need to compute the limit, just from one side, e.g. due to the restrictions in the domain of the function. That is what we call **one-sided limit**. We usually speak about **limit from the right**, denote it by the small plus in the superscript $x \to x_0^+$, and **limit from the left**, denoted analogically by $x \to x_0^-$. We also often speak about left neighborhood $N_-(x_0) = (x_0 - \delta, x_0)$ and right neighborhood $N_+(x_0) = (x_0, x_0 + \delta)$ of theo point x_0.
Strictly speaking, the limits for $x \to \infty$ and $x \to -\infty$, can be also regarded as one sided limits.

Theorem

The function f has the limit L for $x \to x_0$ if and only if

$$\lim_{x \to x_0^-} f(x) = L = \lim_{x \to x_0^+} f(x).$$

The one sided limits are used mostly for the computing the expressions $a/0$:

We say that

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}, \qquad \text{with } \lim_{x \to x_0} f(x) = a \text{ and } \lim_{x \to x_0} g(x) = 0$$

is an expression of the form $a/0$. For its value it holds:

$$\frac{a}{0} = \infty \qquad \begin{cases} a > 0 & \text{and } g(x) \text{ positive} \\ a < 0 & \text{and } g(x) \text{ negative} \end{cases}$$

$$\frac{a}{0} = -\infty \qquad \begin{cases} a > 0 & \text{and } g(x) \text{ negative} \\ a < 0 & \text{and } g(x) \text{ positive} \end{cases}$$

Example 92

Compute the limit

$$\lim_{x \to 0^+} \frac{\sqrt{x}}{x^2}.$$

As $D = (0, \infty)$, we can only compute the limit at $x = 0$ from the right only. The expression is of the form $0/0$ and we can use l'Hôpital rule, as usual:

$$\lim_{x \to 0^+} \frac{\sqrt{x}}{x^2} \overset{\text{l'H}}{=} \lim_{x \to 0^+} \frac{\frac{1}{2}x^{-\frac{1}{2}}}{2x} = \lim_{x \to 0^+} \frac{1}{4x\sqrt{x}} = \text{“}\frac{1}{+0}\text{”} = \infty.$$

Example 93

Compute the limit

$$\lim_{x \to 1} \frac{1}{x^2 - 1}.$$

The limit is of the form $1/0$ and the denominator changes sign at $x = 1$. Therefore, we compute the one-sided limits:

$$\lim_{x \to 1^-} \frac{1}{x^2 - 1} = \frac{1}{0} = -\infty,$$

as the denominator is negative on a left neighborhood, e.g. $N_-(1) = (0, 1)$,

$$\lim_{x \to 1^+} \frac{1}{x^2 - 1} = \frac{1}{0} = +\infty,$$

as the denominator is positive on a right neighborhood, e.g. $N_+(1) = (1, 8)$.
As the one sided limits have different values, the standard limit, $\lim_{x \to 1} \frac{1}{x^2-1}$ does not exist.

L'Hôpital rule and indeterminate expressions

The indeterminate expressions not covered by l'Hôpital rule may be rearranged to the form $\frac{\pm\infty}{\pm\infty}$ or $\frac{0}{0}$.
We start with the expressions of the form $0 \cdot \infty$. For them it holds:

$$0 \cdot \infty = 0 \cdot \frac{1}{\frac{1}{\infty}} = \frac{0}{0} \qquad \text{or} \qquad 0 \cdot \infty = \frac{1}{\frac{1}{0}} \cdot \infty = \frac{\infty}{\infty}.$$

Which one should we use? Generally speaking the one, which provides nicer expression for the derivative.

Example 94

Compute $\lim_{x \to 0^+} x \cdot \ln x$.

The expressions of the form $0 \cdot (-\infty) = (-1) \cdot 0 \cdot \infty$. We show the both ways of tranformation:

$$\lim_{x \to 0^+} x \cdot \ln x = \lim_{x \to 0^+} \frac{1}{\frac{1}{x}} \cdot \ln x = \lim_{x \to 0^+} \frac{\ln x}{\frac{1}{x}} \overset{\text{l'H}}{=} \lim_{x \to 0^+} \frac{\frac{1}{x}}{(-1)x^{-2}} = \lim_{x \to 0^+} -x = 0$$

We just note that the other possibility leads to nowhere.

We turn to the expression $\infty - \infty$. The infinities most often arise when we divide by the variable x anf let x go to zero. The expression $0/0$ can be handled by l'Hôpital rule. For the expressions with non-zero numerator, i.e. $a/0$ we use the procedure prom the previus page, $a/0 = \pm\infty$.
That is why the expression $\infty - \infty$ may be usually expanded with a common denominator, which transforms it into either $0/0$ or $a/0$.

Example 95

Compute $\lim_{x \to 1} \left(\frac{1}{\ln x} - \frac{2}{x - 1} \right)$.

First we rewrite the bracket using the common denominator $(x-1)\ln x$:

$$\lim_{x\to 1}\left(\frac{1}{\ln x}-\frac{1}{x-1}\right)=\lim_{x\to 1}\left(\frac{x-1-\ln x}{(x-1)\ln x}\right),$$

which is of the form $0/0$, and next we can use L'Hôpital rule

$$\begin{aligned}\lim_{x\to 1}\left(\frac{x-1-\ln x}{(x-1)\ln x}\right) &\overset{\text{l'H}}{=} \lim_{x\to 1}\left(\frac{1-\frac{1}{x}}{\ln x+(x-1)\frac{1}{x}}\right)=\left(\frac{0}{0}\right)\\ &\overset{\text{l'H}}{=} \lim_{x\to 1}\left(\frac{-\frac{1}{x^2}}{\frac{1}{x}+\frac{1}{x^2}}\right)=\lim_{x\to 1}\frac{1}{x+1}=\frac{1}{2}\end{aligned}$$

Exercise 96

Compute the following limits:

a) $\lim_{x\to\infty}\left(\sqrt{x^2+2x}-2x\right)$

b) $\lim_{x\to 0} x\cdot\cot(x)$

c) $\lim_{x\to\infty}\left(\sqrt{x^2+x-1}-x\right)$

d) $\lim_{x\to-\infty} x\cdot e^x$

6.6 Tangent lines at improper points: the asymptotes

The asymptotes are special tangent lines that meet the graph at infinity. This is why we need to use limits to decide about asymptotes. We have three types of asymptotes:

Definition

Vertical asymptote $a_V : x = x_0$:

$$\lim_{x\to x_0^{\pm}} f(x)=\pm\infty$$

Horizontal asymptote $A_H : y = a$:

$$\lim_{x\to\infty} f(x)=a$$

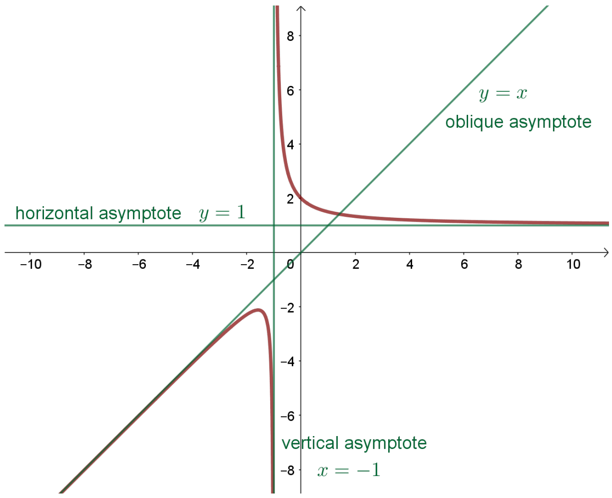
Oblique asymptote $a_O : y = k\cdot x+q$:

$$\begin{aligned}k &= \lim_{x\to\infty}\frac{f(x)}{x}\\ q &= \lim_{x\to\infty} f(x)-k\cdot x\end{aligned}$$

Remark

a) We note that when f is elementary continuous function, the only possibility of a vertical asymptote arises at the border points of the domain intervals $D = (a, b)$. Therefore we require only the existence of one one-sided limit for the existence of the vertical asymptote.

b) In some cases, there might exist a different oblique asymptote for $x \to -\infty$.



The three types of asymptotes.

Example 97

Write down the equations of all asymptotes to the graph of the function

$$f : y = \frac{x^2 - 1}{x}.$$

We first determine the domain of f. Due to the x in the denominator, we have $D = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$.

Therefore, the existence of the only possible vertical asymptote $a_V : x = 0$ will be determined by the limit

$$\lim_{x \to 0} \frac{x^2 - 1}{x} = \text{“}\left(\frac{-1}{0}\right)\text{”} \pm \infty$$

As the one-sided limit from the left is ∞ and from the right $-\infty$, the asymptote indeed exists.

Next, we decide if there is horizontal or oblique asymptote. As

$$\lim_{x \to \infty} \frac{x^2 - 1}{x} \overset{\text{l'H}}{=} \lim_{x \to \infty} \frac{2x}{1} = \infty$$

there is no horizontal asymptote.

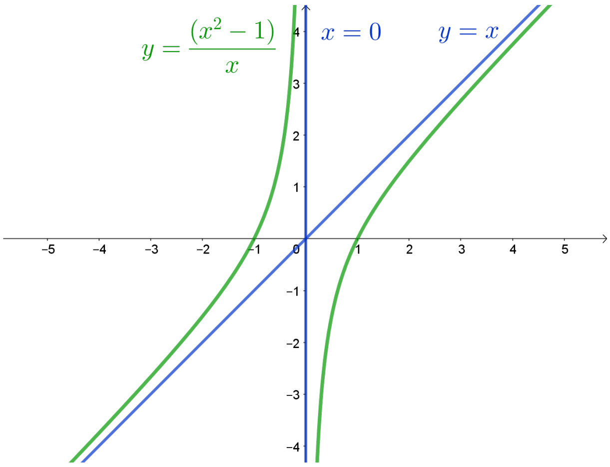
Finally, we compute the formula for the oblique asymptote:

$$k = \lim_{x\to\infty} \frac{\frac{x^2-1}{x}}{x} = \lim_{x\to\infty} \frac{x^2-1}{x^2} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{2x}{2x} = 1$$

$$q = \lim_{x\to\infty} \frac{x^2-1}{x} - 1\cdot x = \text{“}(\infty-\infty)\text{”} = \lim_{x\to\infty} \frac{x^2-1-x^2}{x} = \lim_{x\to\infty} \frac{-1}{x} = 0$$

The same holds also for $x \to -\infty$ and we have unique oblique asymptote $y = x$.

The situation is nicely visible form the graph:



The function $f : y = \dfrac{x^2-1}{x}$ and its asymptotes.

Example 98

Write down the equations of all asymptotes to the graph of the function

$$f : y = \arctan\left(\frac{1}{x}\right).$$

We first determine the domain of f. Due to the $1/x$ in the argument we have $D = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$.

Therefore, the existence of the only possible vertical asymptote $a_V : x = 0$ will be determined by the limit

$$\lim_{x\to 0} \arctan\left(\frac{1}{x}\right) = \text{“}\arctan\left(\tfrac{1}{0}\right) = \arctan(\pm\infty)\text{”}$$

We treat the cases separately. First, we take $x > 0$ and substitute $y = 1/x$. We get

$$\lim_{x\to 0^+} \arctan\left(\frac{1}{x}\right) = \lim_{y\to\infty} \arctan(y) = \frac{\pi}{2},$$

as the function arctan is increasing and $I = (-\frac{\pi}{2}, \frac{\pi}{2})$. Similarly,

$$\lim_{x\to 0^-} \arctan\left(\frac{1}{x}\right) = \lim_{y\to-\infty} \arctan(y) = -\frac{\pi}{2}.$$

Hence, the one-sided limits are different and finite. Therefore, the vertical asymptote does not exist.

Next, we decide if there is horizontal or oblique asymptote. As

$$\lim_{x\to\infty} \arctan\left(\frac{1}{x}\right) = \arctan(0) = 0,$$

we have the horizontal asymptote $y = 0$ and there is no oblique asymptote.

Exercise 99

Draw the asymptotes into the graph of the function $f : y = \arctan\left(\frac{1}{x}\right)$.

Exercise 100

Write down the equations of all asymptotes to the graph of the functions:

a) $y = \dfrac{x-4}{2x+6}$

b) $y = \dfrac{1}{x^2+x-2}$

c) $y = \dfrac{x+1}{x^2-4}$

d) $y = \dfrac{x^2-2x+2}{3x-4}$

e) $y = \dfrac{x^2}{1-x}$

f) $y = 3 - 2x + \dfrac{1}{x^2}$

g) $y = \dfrac{1-x^2}{x^2+3x+4}$

h) $y = \dfrac{x^3+3x^2+1}{x^2+2}$

i) $y = x \cdot \mathrm{e}^{-2x}$

j) $y = x^2 \cdot \mathrm{e}^{-x}$

k) $y = \ln \dfrac{x+1}{x-1}$

l) $y = \dfrac{\sin x}{x}$

m) $y = \dfrac{\cos x}{x}$

n) $y = \dfrac{x}{2} - \cos x$

o) $y = x + \arctan \dfrac{x}{2}$

p) $y = \arctan \dfrac{x+1}{x}$

Name: Workbook for Mathematics I

Department, Institute: Faculty of Mechanical Engineering, Department of Mathematics and Descriptive Geometry

Authors: Jana Bělohlávková, Jan Kotůlek

Place, year of publishing: Ostrava, 2020, 1st Edition

Number of Pages: 82

Published: VSB – Technical University of Ostrava